ISOLATION AND ANALYSIS OF COTTON GENOMIC CLONES ENCOMPASSING A FATTY ACID DESATURASE (*FAD2*) GENE

Wisatre Kongcharoensuntorn, M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2001

APPROVED:

Robert M. Pirtle, Major Professor
Robert C. Benjamin, Committee Member
Kent D. Chapman, Committee Member
Gerard A. O'Donovan, Committee Member
Fritz E. Schwalm, Committee Member
Earl G. Zimmerman, Chair of the Department of Biological Sciences, College of Arts and Science
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Kongcharoensuntorn, Wisatre, Isolation and analysis of cotton genomic clones encompassing a fatty acid desaturase (*FAD2*) gene. Doctor of Philosophy (Molecular Biology), May 2001, 162 pp., 1 table, 30 illustrations, reference list, 101 titles.

Polyunsaturated fatty acids are major structural components of plant chloroplast and endoplasmic reticulum membranes. Two fatty acid desaturases (designated FAD2 and FAD3) desaturate 75% of the fatty acids in the endoplasmic reticulum. The ω-6 fatty acid desaturase (FAD2) may be responsible for cold acclimation response, since polyunsaturated phospholipids are important in helping maintain plant viability at lowered temperatures. To study regulation of *FAD2* gene expression in cotton, a *FAD2* gene was isolated from two genomic libraries using an *Arabidopsis FAD2* hybridization probe and a cotton *FAD2* 5′-flanking region gene-specific probe, respectively. A cotton *FAD2* gene was found to be in two overlapping genomic clones by physical mapping and DNA sequencing. The cloned DNA fragments are identical in size to cotton *FAD2* genomic DNA fragments shown by genomic blot hybridization. The cotton *FAD2* coding region has 1,155 bp with no introns and would encode a putative polypeptide of 384 amino acids. The cotton FAD2 enzyme has a high identity of 75% with other plant FAD2 enzymes. The enzyme has three histidine-rich motifs that are conserved in all plant membrane desaturases. These histidine boxes may be the iron-binding domains for reduction of oxygen during desaturation. To confirm that this FAD2 enzyme is functional, a plasmid construct containing the cotton *FAD2* coding region was transformed into *Saccharomyces cerevisiae*. The transformed yeast cells were able to

catalyze the conversion of oleic acid (C18:1) into linoleic acid (C18:2). The *FAD2* gene contains an intron of 2,967 bp in its 5′-flanking region, 11 bp upstream from the initiation codon. The intron could be essential for transcriptional regulation of *FAD2* gene expression. Several putative promoter elements occur in the 5′-flanking region of this gene. A potential TATA basal promoter element occurs at 41 bp upstream from the cap site. Two presumptive helix-loop-helix (bHLH) motifs that may be seed-specific promoter elements are located at 109 bp and 135 bp upstream from the potential cap site.

ACKNOWLEDGMENTS

This research was supported in part by Grant 99-652 from Cotton Incorporated and University of North Texas Organized Research Funds to Drs. Robert Pirtle and Irma Pirtle.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS..... ii

LIST OF TABLES v

LIST OF ILLUSTRATIONS vi

Chapter

1. INTRODUCTION

I. Fatty acid synthesis in plants ……1
II. Fatty acid desaturases in higher plants… …….7
III. Mutation analysis in plants……..10
IV. Structure of fatty acid desaturasenzymes…….13
V. Possible functions of lipid unsaturation……..…….19
VI. Expression of fatty acid desaturase genes……...….22

2. MATERIALS AND METHODS

I. Bacteriophage libraries, bacterial strains and plasmid DNAs…….28
II. Isolation of a cotton *FAD2* cDNA clone……..29
III. Isolation of cotton *FAD2* genomic and cDNA clones…….30
IV. Large scale phage preparation…….…….33
V. Physical mapping of phage containing putative cotton *FAD2* genes...35
VI. Subcloning of the cotton *FAD2* genomic DNA……...……....36
VII. Generating Nested Deletions……..…….…….43
VIII. DNA sequencing…….…….……45
IX. Genomic DNA gel blot analysis…….…….……46
X. Expression of the cotton *FAD2* gene in yeast cells……..……47

3. RESULTS

I. Screening of the cotton cDNA and genomic DNA libraries……..….54
II. Identification of a cotton ω-6 fatty acid desaturase…….68
III. Amino acid analysis of the *FAD2* coding region……76
IV. Phylogenetic relationships between the cotton FAD2 open reading frame with the open reading frames of other membrane-bound fatty-acyl desaturases…….80
V. Hydropathy plot of the cotton FAD2 amino acid sequence..……..83
VI. Isolation of the cotton *FAD2* gene in the overlapping genomic clone LCFg24……89
VII. Characterization of the overlapping cotton genomic clone LCFg24……..107
VIII. Genomic organization of cotton *FAD2* genes……111
IX. Expression of the cotton *FAD2* coding region in yeast cells…….118

4. DISCUSSION

I. Molecular cloning of the gene encoding a cotton ω-6 fatty acid desaturase (FAD2)……127
II. Analysis of the cotton *FAD2* gene……...129
III. Structural analysis and the topology of the cotton ω–6 fatty acid desaturase (FAD2) polypeptide……..…..133
IV. Expression of the cotton *FAD2* gene in yeast cells…….….138
V. Biotechnology applications with the cotton *FAD2* gene……..142

REFERENCE LIST…….145

LIST OF TABLES

Table Page

1. Homology between the DNA sequence of the cotton *FAD2* gene and other fatty acid desaturase cDNA/gene sequences……...131

LIST OF ILLUSTRATIONS

Figure Page

1. Simplified scheme of plastid fatty acid metabolism…… 2

2. Scheme of the glycerolipid synthesis pathway in leaf cells of *Arabidopsis* 4

3. The structures of acyl-ACP desaturases are composed of di-iron-oxo clusters in a large-helix bundle protein..... 14

4. Topological model of the Δ12 acyl-lipid desaturases in membranes..... 17

5. Agarose gel electrophoresis of putative *FAD2* genomic clones harbored in the phage EMBL3 57

6. Alkaline blot hybridization analysis of *FAD2* genomic clones harbored in the phage EMBL3 59

7. Physical map of the cotton genomic clone LCFg55 harboring a fatty acid desaturase (*FAD2*) gene 62

8. Electrophoresis of restriction endonuclease fragments generated by digestion of the cotton genomic clone LCFg55 DNA on a 1% agarose gel 64

9. Autoradiogram of alkaline blot hybridization of an agarose gel with restriction endonuclease digests of cotton genomic clone LCFg55 DNA 66

10. Agarose gel electrophoresis of deletion plasmid subclones generated from the original construct pCFg55 69

11. Deletion subclone map of the pCFg55 plasmid construct in pDELTA2 71

12. DNA sequence of the coding region of the cotton ω–6 fatty acid desaturase (*FAD2*) gene..... 74

13. Comparison of the deduced amino acid sequences of cotton and other plant FAD2 proteins 77

14. Dendrogram analysis of *FAD*-like DNA sequences in the plant desaturases 81

15. Hydropathy plot of the cotton FAD2 sequence..... 85

Page

16. The amino acid sequence showing the conserved histidine motifs of the cotton FAD2 polypeptide in this dissertation ..... .87

17. Agarose gel electrophoresis of restriction enzyme fragments from the eight genomic clones encompassing the putative *FAD2* gene in the lambda FIXII vector……...91

18. Autoradiogram of an alkaline blot of eight genomic clones encompassing possible *FAD2* genes in the lambda FIXII vector……93

19. Agarose gel electrophoresis of restriction enzyme fragments from the cotton genomic clone LCFg24 DNA on an 0.8% agarose gel……....96

20. The alkaline blot hybridization analysis of LCFg24 DNA fragments 98

21. Agarose gel electrophoresis of restriction fragments derived from digests of lambda LCFg24 DNA on a 0.8% agarose gel..... 100

22. Autoradiogram of the alkaline blot hybridization of an 0.8% agarose gel with restriction enzyme fragments of the cotton genomic clone LCFg24 DNA in Figure 21 102

23. Physical maps of the cotton genomic clone LCFg24 and the related overlapping genome clone LCFg55 104

24. Nucleotide and amino acid sequence of the cotton *FAD2* gene from the clone LCFg24 108

25. Agarose gel electrophoresis of cotton genomic DNA restriction fragments 114

26. Autoradiogram of alkaline blot hybridization of cotton genomic DNA. 116

27. The structure of the plasmid construct designated pYES2/FAD2 transformed into *Saccharomyces cerevisiae* strain INVSc1..... 121

Page

28. Gas chromatographic analysis of the total fatty acid methyl esters extracted from the control *S. cerevisiae* transformed with the pYES2 plasmid 123

29. Gas chromatographic analysis of the total fatty acid methyl esters extracted from *S. cereviasiae* transformed with the plasmid construct pYES2/FAD2..... 125

30. The predicted three-dimensional structure of the putative cotton FAD2 polypeptide (ω-6 fatty acid desaturase) in the endoplasmic reticulum (ER) membrane…….....136

CHAPTER I

INTRODUCTION

I. Fatty acid synthesis in plants

Most of the plant fatty acids found in nature consist of six or seven molecules that have 16 or 18 carbons and one to three double bonds in the acyl group side chains. As shown in Figure 1, these fatty acids are synthesized from acetyl-CoA while attached to an acyl carrier protein (ACP) by a series of reactions for the synthesis of palmitoyl-ACP, stearoyl-ACP and oleoyl-ACP (Ohlrogge, 1994). The introduction of the first double bond occurs while the fatty acid is attached to the ACP. When the fatty acids are released from the fatty acyl-ACPs by the activity of a specific thioesterase, they cross the plastid envelope membrane by an unknown mechanism and are re-esterified to CoA. The desaturation modifications introduce double bonds into fatty acyl chains and then the assembly of three fatty acyl chains is added onto the glycerol backbone to produce triacylglycerols. Subsequently, triacylglycerols are stored in seeds.

As shown in Figure 2, based on the glycerolipid synthesis pathway in leaf cells of higher plants, there are two pathways of fatty acid biosynthesis (Ohlrogge and Browse, 1995). The *de novo* pathway of fatty acid synthesis of 16:0 and 18:1 fatty acids occurs in the chloroplast, and is called the prokaryotic pathway. Both prokaryotic and eukaryotic glycerolipid synthesis are initiated by synthesis of 16:0-ACP and 18:0-

---

The style of this dissertation follows *Plant Cell Physiol*.

Figure 1. Simplified scheme of plastid fatty acid metabolism. The fatty acids are synthesized from acyl-CoAs and the co-substrate, malonyl-ACP, by a series of reactions (3-5). The enzymes represented by the numbers, involved in the pathway, are: 1, acetyl-CoA carboxylase; 2, malonyl-CoA: ACP trans-acylase; 3, 3-ketoacyl-ACP synthase III; 4, 3-ketoacyl-ACP synthase I; 5, 3-ketoacyl-ACP synthase II; 6, stearoyl-ACP desaturase; 7, oleoyl-ACP thioesterase; 8, medium-chain acyl-ACP thioesterase. This diagram was modified from Figure 1 of Ohlrogge (1994).

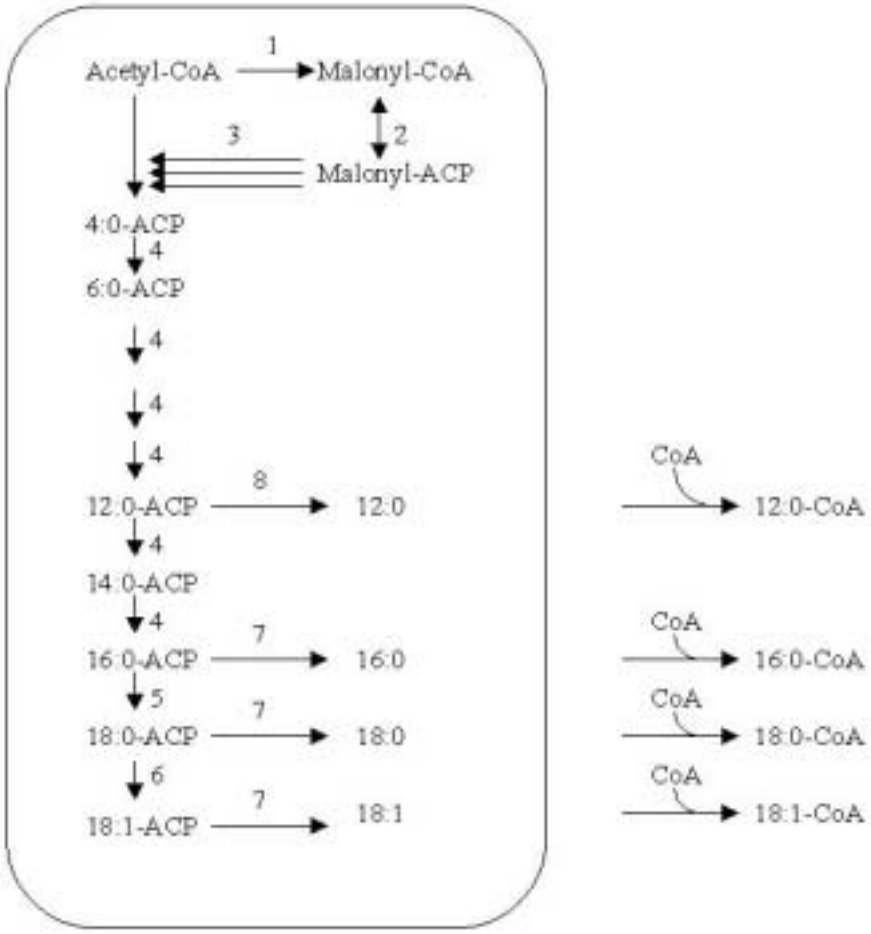
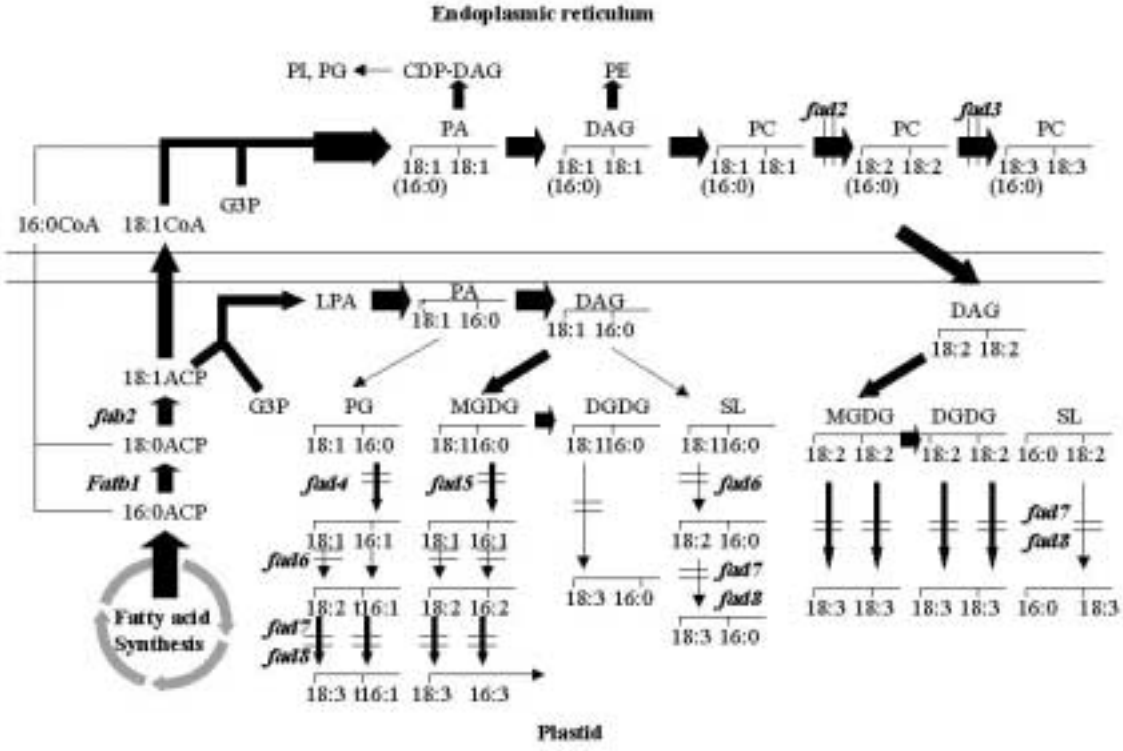

Acetyl-CoA
1
Malonyl-CoA
2
3
Malonyl-ACP
4:0-ACP
4
6:0-ACP
4
4
4
12:0-ACP
8
12:0
4
14:0-ACP
4
16:0-ACP
7
16:0
5
18:0-ACP
7
18:0
6
18:1-ACP
7
18:1
CoA
12:0-CoA
CoA
16:0-CoA
CoA
18:0-CoA
CoA
18:1-CoA

Figure 2. Scheme of the glycerolipid synthesis pathway in leaf cells of *Arabidopsis*. The carbon flux through the various steps is indicated by the broad arrows. The enzymatic deficiencies in various mutants are shown by the broken lines. The series of reactions are divided into two pathways, the prokaryotic pathway in the plastid and the eukaryotic pathway in the endoplasmic reticulum and plastid. Both pathways are initiated by the formation of 16:0-ACP and 18:0-ACP within the plastid. The acyl-ACPs are used within the plastid for phosphatidic acid formation. Phosphatidic acid (PA) is the precursor of mono-and di-galactosyldiacylglycerols (MGDGs and DGDGs), and sulfolipids (SL). In the eukaryotic pathway, the 16:0 and 18:0 fatty acids are converted to their CoA thioesters, and used for synthesis of PA in the endoplasmic reticulum. The PA and C18 fatty acids at the *sn-2* position of PA give rise to phospholipids such as phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylinositol (PI). Desaturation occurs in the endoplasmic reticulum of PC-bound 18:1 to PC-bound 18:2 and 18:3. This diagram is modified from Figure 5 of Ohlrogge and Browse (1995).


Endoplasmic reticulum
PI, PG
CDP-DAG
PE
PA
DAG
PC
fad2
PC
fad3
PC
18:1 18:1 (16:0)
18:1 18:1 (16:0)
18:1 18:1 (16:0)
18:2 18:2 (16:0)
18:3 18:3 (16:0)
16:0CoA
18:1CoA
G3P
LPA
PA
18:1 16:0
DAG
18:1 16:0
DAG
18:2 18:2
18:1ACP
fab2
18:0ACP
fatb1
16:0ACP
G3P
PG
MGDG
DGDG
SL
18:1 16:0
18:1 16:0
18:1 16:0
18:1 16:0
fad4
fad5
fad6
18:1 16:1
18:1 16:1
18:2 16:0
fad6
fad7
fad8
18:2 16:1
18:2 16:2
18:3 16:0
fad7
fad8
18:3 16:0
18:3 t16:1
18:3 16:3
Fatty acid Synthesis
MGDG
DGDG
SL
18:2 18:2
18:2 18:2
16:0 18:2
fad7
fad8
18:3 18:3
18:3 18:3
16:0 18:3
Plastid

ACP within the plastid. In the next step, a double bond is introduced into a 18:0-ACP for conversion into the 18:1-ACP by a soluble desaturase. Most of the 16:0-ACP and the 18:1-ACP is transferred to glycerol-3-phosphate at the *sn-1* and *sn-2* positions, respectively, to form phosphatidic acid (PA). This is called the first step of glycerolipid synthesis in the prokaryotic pathway. A specific phosphatase removes a phosphate group from PA and produces diacylglycerol (DAG). The DAG is a precursor of the major plastid membrane lipids, such as monogalactosyl diacylglycerols (MGDGs), digalactosyl diacylglycerol (DGDGs), and sulfolipids (SLs). By activating the head group of galactose and sulfoquinovose (D-quinovose is 6-deoxyglucose) *via* uridine 5′-diphosphate glucose (UDPG), MGDG, DGDG and SL are synthesized. In contrast, phosphatidyl glycerol (PG), the sole lipid product of the prokaryotic pathway, is synthesized entirely by a cytidine 5′-diphosphate (CDP)-activated form of DAG (CDP-DAG) derived from the reaction of PA with cytidine 5′-triphosphate (CTP).

The first committed step of the eukaryotic pathway starts when 16:0-ACP and 18:1-ACP are hydrolyzed to free fatty acids by thioesterase hydrolysis, and both the 16:0 and 18:1 fatty acids move out of the plastid to enter the endoplasmic reticulum (ER). By the action of an acyl-CoA synthase, the 16:0 and 18:1 fatty acyl-CoA esters are synthesized. These acyl-CoA thioesters are used for synthesis of PA. Thus, the acyltransferases localized in the ER catalyze the modification of the 16:0 fatty acyl chain and 18:1 fatty acyl chain, at the *sn-1* and *sn-2* positions, respectively, to glycerol-3-phosphate (G3P). Subsequently, PA is used to synthesize phosphatidylcholine (PC),

phosphatidylethanolamine (PE) and phosphatidylinositol (PI), which are the characteristic lipid components of extrachloroplast membranes.

It appears that the DAG moiety of PC from the eukaryotic pathway can be returned to the chloroplast envelope (Miquel and Browse, 1992; Browse et al., 1993). Then, these DAG pools are also required for the synthesis of plastid lipids. Moreover, some lipid exchange can occur reversibly between the ER and the chloroplast. There is also evidence that the extra chloroplast membranes in mutants deficient in ER desaturases contain polyunsaturated fatty acids derived from chloroplasts (Miquel and Browse, 1992; Browse et al., 1993).

The major difference between the prokaryotic and eukaryotic pathway is that the glycerolipids derived from the prokaryotic pathway contain 16:0 fatty acyl chains only at position *sn-2*, whereas the glycerolipids from the eukaryotic pathway have 18:1 fatty acyl chains at both the *sn-1* and *sn-2* positions. Also, the prokaryotic pathway of glycerolipid synthesis takes place in plastids, but the eukaryotic pathway is located at extrachloroplast sites, especially in ER and nonphotosynthetic tissues such as oil seeds. Both pathways can function in photosynthetic tissue. However, the eukaryotic pathway is found solely in non-photosynthetic tissues such as oil seeds (Tocher et al., 1998).

II. Fatty acid desaturases in higher plants

Fatty acid desaturases are enzymes that catalyze the formation of double bonds in fatty acids. There are two major types of desaturases in plants, a soluble chloroplast)9- desaturase and membrane-bound desaturases that desaturate the fatty acyl groups of phosphatidylcholine to produce Δ12- and Δ15- fatty acids, as well as several other

desaturated fatty acids (shown in Figure 2). The major glycerolipids in all plant tissues are desaturated by two different types of desaturases, soluble and membrane-bound desaturases, which occur in chloroplast and ER, respectively.

The first type of desaturase is the Δ9 (stearoyl) desaturase which introduces a *cis*-double bond at the Δ9 position of stearoyl-ACP to produce oleoyl-ACP. This enzyme is responsible for the conversion of saturated fatty acids to unsaturated fatty acids in vegetable oils. The soluble chloroplast desaturase needs ferredoxin as the electron donor, and it has a conserved region at the C-terminal end. The first Δ9 (stearyol-ACP) desaturase gene was isolated from safflower, and the polypeptide gene product is a soluble dimer of a 38-kDa subunit which contains a 33-amino acid transit peptide (Slabas and Fawcett, 1992).

Membrane-bound desaturases, such as the Δ12- and Δ15-desaturases, convert oleate into linoleate and linolenate, respectively. Since these desaturases are integral membrane proteins in the endoplasmic reticulum, they are quite difficult to isolate and characterize by biochemical methods. However, most information about the number and properties of these plant desaturases has been obtained by isolation and characterization of a series of mutants involved in lipid biosynthesis in *Arabidopsis thaliana.* There are four chloroplast desaturase loci: *FADA, FADB, FADC* and *FADD,* which have been renamed *FAD4, FAD5, FAD6*, and *FAD7,* respectively (Ohlrogge and Browse, 1995).

The *FAD4* gene products, called Δ-3 desaturases, insert a *trans* double bond into 16:0-fatty acyl chains at the *sn-2* position of PG. The polypeptide product of the *FAD5*

gene acts on the synthesis of Δ7- 16:1 fatty acyl chains in MGDGs and DGDGs. It is different from the other two chloroplast desaturases which function on acyl chains with no specific fatty acyl chain (16:0 or 18:0) at the *sn-1* or *sn-2* positions of the glycerol backbone. In addition, the C16:1/18:1 desaturase is the product of the *FAD6* gene, whereas C16:2/18:2 desaturases are encoded by the *FAD7* and *FAD8* genes. Two additional membrane-bound desaturase genes are the *FAD2* and *FAD3* genes, which encode the C18:1(FAD2) and C18:2(FAD3) desaturases found in the ER. The FAD2 and FAD3 enzymes, also called ω-6 desaturase and ω-3 desaturase, respectively, function on fatty acyl chains at both the *sn-1* and *sn-2* positions of PC (Harwood, 1997). The DNA sequence analysis of these desaturase genes indicates that there are a number of His-containing conserved motifs in the corresponding expressed polypeptide products (Shanklin et al., 1994; Shanklin and Cahoon, 1998). The ER ω-3 desaturases have been isolated from *Glycine max*, *Brassica napus*, *Vigna radiata*, and *Nicotiana tabacum* (Tocher et al., 1998). Most of the presumptive membrane-bound desaturase polypeptides (such as FAD3) have a group of eight conserved histidine residues, composed of a tripartite motif $HX_{(3-4)}HX_{(7-41)}HX_{(2-3)}HHX_{(61-189)}HX_{(2-3)}HH$, that is homologous to that of the Δ12-desaturase from cyanobacteria (Shanklin and Cahoon, 1998). Plastid ω-3 desaturase genes have been isolated from *G. max*, *B. napus, Ricinus communis*, and *Limnanthes douglasis* (Kodama et al., 1997). Moreover, the ER ω-6 desaturase gene (*FAD2*) has been cloned in *Arabidopsis thaliana* (Okuley et al., 1994) and the cDNA in *Gossypium hirsutum* (Liu et al., 1999)

III. Mutation analysis in plants

Fatty acid desaturases can be studied by isolating plants with various mutations in the fatty acid desaturase genes and monitoring the proportions of various fatty acids observed in different plants. The plant model most widely used for characterizing the desaturase genes is *Arabidopsis*. At least 12 of the reactions and seven types of mutants of *Arabidopsis* correspond to enzymes of the major pathways of glycerolipid synthesis in leaf cells (Browse, 1991). Also, mutants that pinpoint the effects of desaturases are involved in leaf glycerolipid synthesis. Most desaturase mutants fall into two groups, based on the loss or reduction of seed lipid components. All of these desaturase mutants have been observed in plants such as safflower, maize, flax, and soybean. However, it is not possible to describe all mutants that have variations in fatty acid composition. Thus, only selected examples from the recent literature, most of which describe mutants affecting seed and leaf lipids, are mentioned (Ohlrogge and Browse; 1995; Harwood, 1997).

Mutations affecting seed lipid composition

The most interesting examples of the mutagenic approach are two ethylmethane sulfonate-induced mutants of soybean. These types of mutants, called low-linolenate mutants, contain less than one percent of the wild-type level of C18:3 $^{\Delta 9,12,15}$ and an increase in 18:2 $^{\Delta 9,12}$. These low-linolenate mutants are due to a co-dominant nuclear mutation in the *FAD2* locus, so it lacks the C18:2 $^{\Delta 9,12}$-desaturase activity in the eukaryotic pathway (Ohlrogge et al., 1991). Also, high oleic mutants have a defect in

two dominant alleles of the desaturase gene, *ol* and *ml*, such that the mutants have lost C18:1 $^{\Delta 9}$ desaturase activity (Ohlrogge et al., 1991).

In *Arabidopsis*, mutants deficient in oleate desaturation were reported to have low seed lipid composition in C18: 2 $^{\Delta 9,12}$ and C18:3 $^{\Delta 9,12,15}$ fatty acids, but a high seed composition in the 18:1$^{\Delta 9}$ fatty acid. This is due to a single nuclear recessive mutation at the *FAD2* locus, such that the *FAD2* mutants lack C18:1 $^{\Delta 9}$ desaturase activity (Ohlrogge et al., 1991). The *Arabidopsis* mutants have increased levels of C18: $^{\Delta 9,12}$ and decreased levels of C18:3 $^{\Delta 9,12,15}$. This type of mutant is caused by a mutation at the *FAD3* locus. These mutants are the result of a single co-dominant nuclear mutation of the *FAD3* locus. The *FAD3* mutant is expressed not only in seeds, but in the leaves and roots as well (Ohlrogge et al., 1991).

Mutations affecting leaf lipid composition

Most studies of mutants with alterations in leaf lipid composition have been done in *Arabidopsis*. By determining the reduction of unsaturated fatty acids and the accumulation of unsaturated precursor molecules, *Arabidopsis* desaturase mutants can be divided into four groups: *FAD4, FAD5, FAD6* and *FAD7* mutants (Ohlrogge et al., 1991). These kinds of mutants are supposed to have a single, recessive nuclear mutation (Ohlrogge et al., 1991). However, there is no mutant that has a defect at the stearoyl-ACP desaturase locus. A single recessive nuclear mutation at the *FAD4* locus results in the absence of C16:1 trans-Δ3 and an increase in C16:0. This mutant acts specifically on C16:0 at the *sn*-2 position of phosphatidylglycerol (PG) (Ohlrogge et al., 1991).

The *FAD5* mutant has reduced amounts of *cis*-unsaturated C16 carbon fatty acids and an accumulation of palmitic acid. Compared to the *FAD4* mutant, the *FAD5* desaturase mutant affects the desaturation of C16:0 fatty acid to C16:1 at the *sn-2* position of monogalactosyldiacylglycerol (MDGD) instead of at PG (Ohlrogge et al., 1991). Also, indications are that the *FAD5* mutant significantly affects the proportion of chloroplast lipids synthesized by the prokaryotic pathway (Somerville and Browse, 1991).

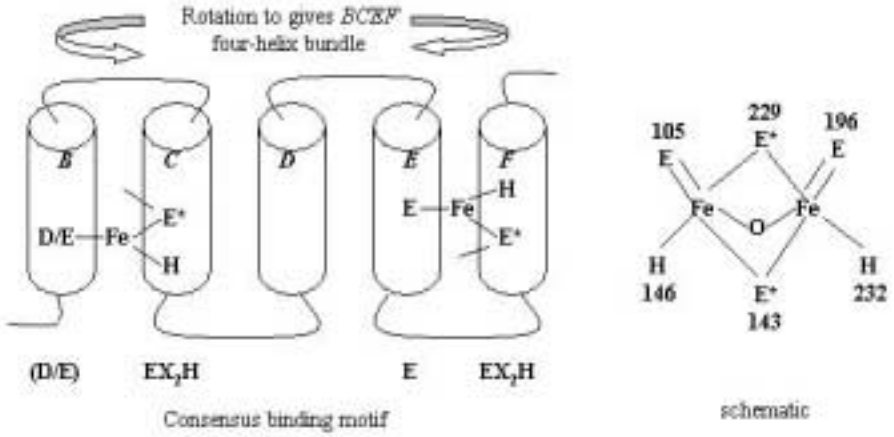
One of the mutants in the *FAD6* locus was defective in both C16 and C18 polyunsaturated fatty acids, and had corresponding accumulations of C16:$1^{\Delta 7}$ and C18:$1^{\Delta 9}$. Both the *sn-1* and *sn-2* positions of all the major chloroplast lipids (MGDG, DGDG, SQD, and PG) were affected. However, *FAD7* mutants had no significant effects on overall fatty acid composition in the acyl chain or head group on the glycerol (Ohlrogge et al., 1991). Studies of the *FAD7* locus mutants suggested that *FAD7* mutants had significant decreases in the levels of both C16:3 $^{\Delta 7,10,13}$ and C18:3 $^{\Delta 9,12,15}$ fatty acids in extracts of whole leaves and corresponding increases in the amounts of C16:2 $^{\Delta 7,10}$ and C18:2 $^{\Delta 9,12}$ fatty acids. The effect on chloroplast lipids in the *FAD7* mutant occurs at the *sn-1* and *sn-2* positions of MDGD and DGDG, and the *sn-1* position of PG and sulfolipids (SL) (Ohlrogge et al, 1991). Also, the *FAD7* mutation is expressed at temperatures above 28°C. This type of mutant is supposed to be a cold-induced desaturase mutant. Thus, the effects of all loci of the various *FAD* genes can be used to interpret the regulation of membrane lipid desaturation.

IV. Structure of fatty acid desaturase enzymes

Fatty acid desaturase enzymes can be divided into three classes: acyl-ACP, acyl-CoA, and acyl-lipid desaturases (Shanklin and Cahoon, 1998; Los and Murata, 1998; Tocher et al., 1998). The acyl-ACP desaturases found in the stroma of plants are the desaturases that introduce double bonds into fatty acids esterified to acyl carrier protein (ACP). The acyl-CoA desaturases are found in animals, yeast and fungi, and are membrane-bound enzymes associated with the ER. These desaturases introduce unsaturated bonds into the CoA esters of the fatty acid (Los and Murata, 1998). The other desaturase class contains the acyl-lipid desaturases, for desaturation of fatty acids esterified in glycerolipids. The acyl-lipid desaturases are also membrane-bound enzymes associated with the ER, the plant chloroplast membrane, and cyanobacterial thylakoid membranes (Tocher et al., 1998, Los and Murata, 1998).

Soluble desaturases, such as acyl-ACP desaturases, are commonly found in plants. The best example is the Δ9 acyl-ACP desaturase which catalyzes desaturation of stearic acid to oleic acid in the stroma of chloroplasts. The acyl-ACP desaturase enzyme is a homodimer with a diiron complex at the center (Figure 3). The two iron atoms interact with O_2 with a consensus-binding motif of $[(D/E)X_2H_2]$ (Shanklin and Edgar, 1998; Los and Murata,1998). The crystallographic analysis of stearoyl-ACP desaturase from castor seeds (*Ricinus communis*) indicates that this desaturase forms a di-iron-oxo active center, with the two iron atoms bound in the symmetric structure. One of the iron atoms interacts with the side chains of E19 and H232 and the other interacts with the side chains of E105 and H146 (Lindquist et al., 1996). The deep

Figure 3. The structures of acyl-ACP desaturases are composed of di-iron-oxo clusters in a large-helix bundle protein. The left panel shows the organization of the helix structure. The right panel shows the schematic crystallography of the castor Δ9-18:0 ACP desaturase ligation sphere. Glutamates coordinate both of the two irons, as indicated by asterisks (*). This diagram was modified from Figure 2 of Shanklin and Cahoon (1998).

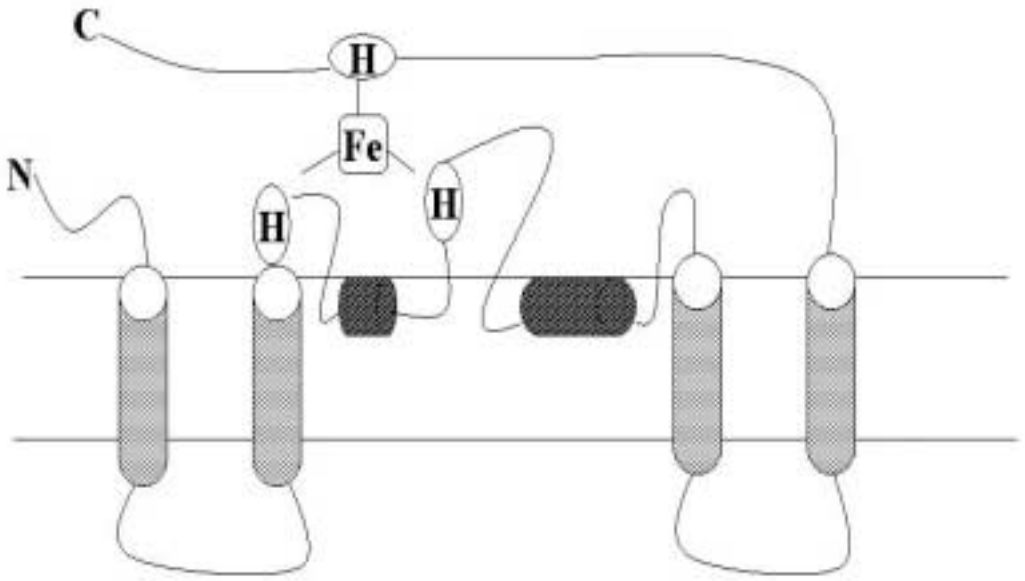

Rotation to gives BCEF
four-helix bundle
B
C
D
E
F
D/E
Fe
E*
H
E
Fe
H
E*
(D/E)
EX2H
E
EX2H
Consensus binding motif
105
E
229
E*
196
E
Fe
O
Fe
H
146
E*
143
H
232
schematic

channel of an extended surface may be the site to which the fatty acyl chain binds.

One of the classes of integral membrane desaturases is the acyl-CoA desaturases, most of which consist of 300-500 amino acid residues and span the membrane lipid bilayer four times (Fox et al., 1993; Shanklin et al., 1994). These desaturases, such as the Δ5, Δ6 and Δ9 acyl-CoA desaturases of animal, yeast, and fungi, are the electron acceptors of the electron-transport system that contain cytochrome b_5 and a NADH-dependent cytochrome b_5 reductase (Mitchell and Martin, 1995; Jeffcoat et al., 1977). The cDNAs for the stearoyl CoA desaturases have been cloned and their corresponding polypeptides have unique structures with the three histidine-rich conserved motifs (Meesters et al., 1997; Mizutani et al., 1998; Gibson, 1994).

The other class of integral membrane desaturases are the acyl-lipid desaturases. These enzymes, found in cyanobacterial cells and chloroplasts, introduce a double bond into the fatty acyl chain of polar glycerolipids. The structure of these desaturases is similar to the acyl-CoA desaturases, which are transmembrane proteins spanning the membrane four times with 300-500 amino acid residues (Murata and Wada, 1995) as shown in Figure 4. The cyanobacterial desaturase uses ferredoxin as the electron donor. In contrast, the desaturases in the cytoplasm of plant cells require a system that is composed of cytochrome b_5 and a NADH:cytochrome b_5 oxidoreductase (Kearns et al., 1991). The cDNAs for acyl-lipid specific isoforms of desaturases have been cloned for the cyanobacterium Synechocystis sp. PCC6803. The Δ6, Δ9, Δ12 and ω-3-acyl-lipid desaturases have been characterized and are specified by the double bond that is

Figure 4. Topological model of the Δ12 acyl-lipid desaturases in membranes. The conserved histidine clusters represented by the circled H's and iron atoms are located in the catalytic center on the cytoplasmic side. The enzyme is proposed to be a transmembrane protein with four membrane-spanning α-helices represented by the stippled cylinders. This diagram was modified from Figure 2 of Los and Murata (1998).



The predicted structure of the Δ 12 acyl-lipid desaturase in the membrane

introduced nearest the carboxyl terminus (Δ-position) or the methyl terminus (ω-position) of the fatty acid (Reddy et al., 1993; Sakamoto et al., 1994b; Sakamoto et al., 1997; Murata et al., 1995). It is assumed that histidine clusters and iron ions are the components of the catalytic center of the desaturases. Site-directed mutagenesis of the Δ-12 acyl-lipid desaturase gene of *Synechocystis* sp. PCC6803 showed that substitution of any of the conserved histidine residues by another amino acid caused the loss of enzymatic activity (Avelange-Macherel et al., 1995; Schneider et al., 1992).

V. Possible functions of lipid unsaturation

There is much evidence indicating the possible roles of fatty acid unsaturation in membrane structure and function in plants, such as changes in the ultrastructure of the chloroplast membrane, the effect on thermal tolerance, the action of signaling molecules such as jasmonic acid, and pathogen response (Hamada et al., 1996; Farmer, 1994; Somerville and Browse, 1991; Kirsch et al., 1997). To study the roles of trienoic fatty acids on chloroplast structure and photosynthetic function, the *FAD7* mutant of *Arabidopsis* was used. A large reduction of C18:3 and C16:3 acyl groups in membrane lipids caused a 45 percent reduction in the cross-sectional area of chloroplasts, and correspondingly decreased the amount and distribution of lamellar membranes as shown in electron micrographs (Kodama et al., 1994). Thus, the reduction in chloroplast size is compensated for by an increase in 45% of the chloroplasts per cell in the mutants. However, the large change in the number of trienoic acids had only a minor effect on sizes of the chloroplasts at room temperature (De Lorgeril et al., 1996). This report implied that the frequent division of chloroplasts is controlled by changes in unsaturated

fatty acid content. Very slight changes in the degree of unsaturation was adequate to cause the MGDGs to undergo a lipid-to-gel phase transition at room temperature (Gounaris et al., 1983), meaning that the unsaturated fatty acid content is the major factor that maintains fluidity of the plant cells.

Fatty acid desaturases may have a role during high and low temperature growth in soybean tissue cultures during seed development. Cultures of seedpods were grown for 20 hours at 20°C, 25°C, or 35°C. The linolenyl- and oleoyl-desaturase activities in seeds, cultured at 20˚C and 25˚C respectively, were found to be 94 and 10 times as active, relative to the expression controls. In contrast, the activities of both desaturases were diminished in seeds cultured at 35°C. These results suggest that the enzymes for fatty acid desaturation can be modulated to respond to alterations in growth temperature, whereas the enzymes for fatty acid synthesis and elongation are not affected by changes in growth temperature (Cheesbrough, 1989).

That increased amounts of trienoic fatty acids enhance cold tolerance was reported in transgenic tobacco (Kodama et al., 1994). The *FAD7* desaturase gene was introduced into transgenic tobacco plants, and the proportions of various fatty acids was determined. The results indicated that the increased production of the trienoic fatty acids hexadecatrienoic (C16:3) and linolenic (C18:3) occurred in response to cold temperatures, protecting the transgenic plants against cold damage. The *FAD7* gene was introduced into transgenic tobacco cells, the cells were exposed to 1°C for seven days, and then they were cultured at 25°C. The results showed that transgenic tobacco plants overexpressing the *FAD7* gene had a seven percent increase in trienoic fatty acids

during cold treatment as compared to cold-treated wild type plants. Therefore, increasing amounts of trienoic fatty acids by introducing the *FAD7* desaturase gene may alleviate cold damage (Okuley et al., 1994).

Similar to the effects of cold acclimation in plants, transgenic cyanobacteria, such as *Anacystis nidulans*, also have tolerance to cold temperatures. An experiment was done by cloning the *desA* gene, a plant desaturase gene, and introducing the gene into *Anacystis nidulans*, which lacks C18:2 fatty acids. The results indicate that appreciable amounts of C18:2 fatty acids were present (six percent) in *A. nidulans* when exposed to low non-freezing temperatures. Therefore, chilling sensitive cyanobacteria such as *A. nidulans* apparently increases their tolerance to low temperature (Wada et al., 1990).

Cold-sensitive organisms, such as the cyanobacterium *Synechocystis* PCC6803, also have been studied for the changes in lipid composition. Based on screening a *FAD2* mutant which is not able to grow at low temperatures, the mutant was shown to lack a Δ12 desaturase (Wada and Murata, 1989). *FAD2* and *FAD6* mutants of *Arabidopsis* have also been identified that have defects in specific isoenzymes of the plant desaturases equivalent to the enzyme encoded by the *FAD2* gene of *Synechocystis* PCC6803.

Trienoic fatty acids are not only important in low temperature tolerance, but also are important in signal transduction in plants. Traumatic acid (12-oxo-(10ε)-dodecenoic acid) and jasmonic acid (octadecanoid-derived cyclopentanones and cyclopentanols) are examples of precursors of fatty acid-signaling molecules which are involved in the

defensive response of higher plant cells. Hamada et al. (1996) isolated a cDNA encoding a plastid ω-3 fatty acid desaturase from tobacco, and tested the effects of wounding on expression of this cDNA in transgenic tobacco plants. The results indicate that expression of the ω-3 desaturase increased in response to wounding the tobacco leaf tissue. Thus, wounding enhances the conversion of linoleic acid to linolenic acid in leaf polar lipids (Hamada et al., 1996).

Fatty acid desaturation and transcription of *FAD2* genes may play a role in the pathogen response in plants. The ER Δ12 desaturase in parsley (*Petroselinium crispum*) is transiently induced by fungal infection (Kirsch et al., 1997). Treatment of cultured parsley cells with the fungal peptide elicitor Pep25 activated rapid and large changes in polyunsaturated fatty acid composition. The Δ12 fatty acid desaturase mRNA accumulated rapidly and transiently in protoplasts and leaves of the elicitor-treated parsley cells.

VI. Expression of fatty acid desaturase genes

Okuley et al. (1994) cloned the gene for *Arabidopsis FAD2* by T-DNA insertional mutagenesis. *FAD2* mutants were chosen and used to produce T-DNA insertions containing *FAD2* mutations. The T-DNA was used as a probe in order to isolate *FAD2* cDNA and genomic clones from *Arabidopsis* cDNA and genomic DNA libraries. The size of the *FAD2* gene was three kilobases (kb), and the deduced amino acid sequence of the expressed polypeptide has strong identity with other plant desaturases. By Northern blot analysis, the *FAD2* mRNA was shown to be expressed throughout the plant and the transcription levels of the *FAD2* gene are increased

commensurately for oleate desaturation. Interestingly, the expression of the *Arabidopsis FAD2* gene in the yeast *S. cerevisiae* under control of the *GAL1* promoter indicated that the C18:2 fatty acids were increased, but the C16:0 fatty acids were decreased when cultured at 8°C, and that the tolerance of the cells to 15% ethanol was also increased (Kajiwara et al., 1996).

Gibson et al. (1994) isolated an *Arabidopsis* gene encoding a ω-3 fatty acid desaturase (*FAD8*) by using a partial sequence of the *Arabidopsis FAD3* gene as a probe. The *FAD8* cDNAs were isolated from cDNA libraries prepared from plants grown at ambient temperature. The *FAD8* cDNA is 1,308 bp in length with an open reading frame of 435 amino acids (GenBank Accession No. U08216). As compared to the *FAD7* gene products, the *FAD8* polypeptide would function equivalently. The two genes have 75% nucleotide identity. The *FAD8* seems to be a temperature-sensitive gene, whereas the *FAD7* gene is not.

Expression of the *Arabidopsis FAD7* gene and its regulation was studied in transgenic tobacco plants by fusion of the *FAD7* promoter to the β-glucuronidase (*GUS*) or the luciferase (*LUC*) reporter genes in tobacco plants (Nishiuchi et al., 1995). A deletion analysis of the 5′-region of the *FAD7* promoter was done in order to ascertain which regions of the *FAD7* promoter regulated expression of the β-glucuronidase and the luciferase reporter genes. The results showed that 825 base pairs (bp) of the *FAD7* promoter fragment upstream from the transcription start-point contained *cis*-acting elements such as BoxII (GT-1 sites) and G-box-like (CCACTTGG) motifs. Both GT-1 sites and G-boxes are members of light-responsive promoters (Guilfoyle, 1997).

In soybean, *FAD2-1* and *FAD2-2* genes encoding ω-6 desaturase have also been isolated and characterized. First, a *FAD2-1* cDNA clone was isolated from a developing soybean embryo library. Using the *FAD2-1* cDNA as a probe, a *FAD2-2* genomic clone was isolated from a soybean genomic DNA library. The functions of both *FAD2-1* and *FAD2-2* genes were investigated by complementation with *FAD2-1* mutants. The results indicated that the soybean *FAD2-2* gene was expressed in both vegetative tissues and developing seeds, whereas the *FAD2-1* gene was expressed in developing seeds. However, the *FAD2-1* and *FAD2-2* genes were not up-regulated by exposure to low temperature (Heppard et al., 1996).

Although Δ6- and Δ5-fatty acyl desaturases are hard to identify in higher plants, Δ6-desaturase activity was demonstrated in the microsome fraction from *Borago officinalis*. PCR-based cloning was used to isolate the borage microsomal Δ6 desaturase gene. The degenerate primers were designed from highly conserved regions of the desaturase sequence and a cDNA library generated from total RNA of developing cotyledons of borage was used as template DNA. A full-length cDNA clone (pBdes 6) was isolated, and found to have 1,344 bp with a coding region for 448 amino acids. The expression of the pBdes6 clone was studied by using the Ω-enhanced cauliflower mosaic virus 35S promoter by *Agrobacterium*-mediated gene transfer, and the pBdes6 gene was introduced into tobacco plants. The transgenic tobacco plants showed the accumulation of 18:3 $^{\Delta 6,9,15}$ and 18:4 $^{\Delta 6,9,12,15}$ fatty acids to levels of 13.2% and 9.6% respectively within leaves. In contrast to other soluble desaturases, the borage

microsomal Δ6- desaturase has an N-terminal domain with cytochrome b_5 activity (Sayanora et al., 1997).

Schmidt et al. (1994) cloned a membrane-bound plastidial ω-6 desaturase cDNA from spinach by PCR. First, they purified a 40 kDa ω-6 desaturase by anion exchange, cation exchange, and ferredoxin-affinity column chromatography. Eight of 13 N-terminal amino acids were used to design a degenerate inosine-containing oligonucleotide primer for PCR amplification. The 5′ end of the cDNA was amplified by 5′ Rapid Amplification of cDNA Ends (5′-RACE). Sequences from both 5′-RACE and 3′-RACE were used to construct primers, and the full-length cDNA clone was amplified by PCR. The sequence of the ω-6 desaturase cDNA was characterized, and the deduced amino acid sequence has about 50% identity with the cyanobacterial ω-6 desaturase. The spinach ω-6 desaturase contains the three histidine boxes highly conserved among membrane-bound desaturases (Schmidt et al., 1994).

Hitz et al. (1994) cloned ω-6 desaturases from *Glycine max* and *Brassica napus*, using a novel probes derived from conserved amino acid sequences between plant and cyanobacterium ω-6 fatty acid desaturases. Oligonucleotides derived from the conserved amino acid sequences were used as probes to isolate the clone from cDNA libraries of *G. max* and *B. napus*. Both clones from these two cDNA libraries were sequenced and the amino acid sequences deduced. The expression of the *B. napus* cDNA was tested by transforming the *B. napus* chimeric gene into the cyanobacterium *Synechococcus* sp. PCC7942, which lacks polyunsaturated fatty acids. The $16{:}1^{\Delta 9}$ fatty acids were converted to $16{:}2^{\Delta 9,12}$ fatty acids in the transformed cells.

Many fatty acid desaturases corresponding to the *Arabidopsis FAD* genes have been isolated and their expression studied by genetic engineering. For example, the *FAD3* gene has been isolated by both T-DNA tagging (Yadav et al., 1993) and map-based cloning (Arondel et al., 1992). The *FAD2* gene was also isolated by T-DNA tagging (Okuley et al., 1994). Chromosome walking with artificial yeast chromosomes was used to isolate a clone for the *FAD7* gene (Iba et al., 1993). The *FAD8* gene was identified by using the *FAD3* gene as a hybridization probe (Gibson et. al., 1994). In addition, several other plant *FAD2, FAD3, and FAD7* genes have been isolated by heterologous probing with the *Arabidopsis* or *Brassica* genes (Yadav et al., 1993; Hamada et al., 1996a; Okuley et al., 1994; Van de Loo et al., 1994). Other methods (such as PCR) were used to purify a ω-6-desaturase cDNA from spinach by using 5′ RACE and 3′ RACE as primers (Schmidt et al., 1994).

The regulation of expression of plant fatty acid desaturases has not yet been thoroughly studied. However, the research described in this dissertation should provide information on the structure and regulation of the cotton *FAD2* gene and begin to delineate the role of fatty acid desaturase gene expression in membrane fatty acid composition.

Thus, the purpose of this research was to isolate a ω–6 desaturase gene (*FAD2*) from *Gossypium hirsutum* in order to determine its structure and regulation of expression. A cotton *FAD2* gene has been isolated by screening cotton genomic DNA libraries. The gene structure has been characterized by physical mapping and DNA sequencing. The gene copy number of the *FAD2* gene in the cotton genome has been

determined by genomic blot hybridization. In addition, the cotton *FAD2* coding region has been functionally expressed in transformed yeast cells.

The regulation of expression of the *FAD2* gene may be important in understanding the composition of fatty acids in plant membrane phospholipids. This is of crucial importance in understanding the basis of membrane fluidity in the cold tolerance of plants.

CHAPTER II

MATERIALS AND METHODS

I. Bacteriophage libraries, bacterial strains and plasmid DNAs

A cotton cDNA library constructed from cotyledon mRNA of 48-hour dark grown cotton seedings (*Gossypium hirsutum* cv. Deltapine 62) and harbored in the Stratagene UniZAP lambda vector was kindly provided by Dr. R. N. Trelease of Arizona State University (Ni and Trelease, 1991). One cotton genomic library (*Gossypium hirsutum*, cv. Acala SJ-5) in the lambda vector EMBL3 was obtained from Dr. David M. Anderson of Phytogen Seeds, Placentia, CA (Grula et al., 1995). A second cotton genomic library (constructed from *Gossypium hirsutum* L. cv. Acala SJ-2) harbored in the Lambda FIXII vector (Stratagene) was obtained from Dr. Thea Wilkins of the University of California at Davis. The pDELTA2 cosmid vector and *E. coli* strain *DF1* and *E. coli* strain DH10B competent cells (GIBCO BRL Life Technology) were used to construct deletion subclones of one putative *FAD2* genomic clone. The recombinant plasmids containing the 5′-flanking region of the cotton *FAD2* gene in the vector pGEM7Zf (+) (Promega) and the cotton *FAD2* coding region in the yeast-bacterial shuttle vector pYES2 were constructed by Mr. Mongkol Nampaisansuk of our laboratory. For yeast expression, the pYES2 plasmid vector and the *Saccharomyces cerevisiae* strain INVScI were purchased from Invitrogen (San Diego, CA).

II. Isolation of a cotton *FAD2* cDNA clone

The cotton cDNA library was screened using as probe a heterologous 0.5 kb *Sal*I fragment isolated from an *Arabidopsis* cDNA clone designated *FAD2-43* (GenBank T13887, Clone 43A5T7). This clone was provided by Dr. Thomas Newman of the *Arabidopsis* Biological Resource Center, Ohio State University (Newman et al., 1994; Okuley et al., 1994). The cDNA library was screened by the plaque hybridization procedure of Benton and Davis (1977), using ^{32}P-labeled DNA probe generated from the *Arabidopsis* cDNA fragment by the random priming procedure of Feinberg and Vogelstein (1983).

The phage library was amplified in the lambda-sensitive strain *E.coli* pBlueScript P2 from Stratagene overnight at 37°C. Approximately 1 x 10^7 plaques were screened in order to obtain a representative number of the cotton cDNA clones. Therefore, 36 plates at a plaque density of 3 x 10^5 plaque-forming units (pfu) were prepared by mixing the diluted phage with *E. coli* pBluescript. Four ml of top agar was poured onto prewarmed NZY agar plates (0.5% yeast extract, 1% NZY Amine (casein hydrolysate), 0.5% NaCl, 0.02 M Mg_2SO_4). The 36 plates were incubated at 37°C overnight until the plaques were clearly visible and reached a diameter of about 1.5 mm (Benton and Davis, 1977).

For plaque hybridization, positively-charged nylon membranes (Amersham Hybond N^+) were overlayed on NZY plates containing the bacteriophage plaques. India ink was punched through the filter with a syringe needle at arbitrary locations on the filter as orientation markers. The nylon membrane filter replicas were removed from

the NZY plates and transferred successively to denaturing solution (1.5 M NaCl, 0.5 M NaOH), neutralyzing solution (1.5 M NaCl, 0.1 M Tris-HCl, pH 7.3), and rinsing solution (2xSSC, 0.2 M Tris-HCl, pH 7.5) for 5 min, 4 min, and 4 min, respectively. 1xSSC is 0.15 M NaCl and 0.015 M sodium citrate (pH 7.0). Prehybridization of the nylon membranes was performed at 55°C in a solution containing 6xSSC, 0.5% SDS, 10X Denhardt's reagent, 0.2% Ficoll (Type 400, Pharmacia), 0.2% polyvinylpyrrolidone, 0.2% bovine serum albumin (Fraction V; Sigma) and sheared salmon sperm DNA (100 μg/ml) for 4 h. Hybridization was done in a solution of 6xSSC, 0.5% SDS, 5xDenhardt's reagent, 20 mM Tris-HCl (pH 8.0), 2 mM EDTA, 2.5 mM Na pyrophosphate (pH 8.0), denatured sheared salmon sperm DNA (100 μg/ml), and [^{32}P]-labeled probe for approximately 18 h. After hybridization, the nylon membranes were washed once for five min. and once for thirty min. in 2xSSC, 0.1% SDS, and then twice for 30 min. in 1xSSC, 0.1% SDS.

III. Isolation of cotton *FAD2* genomic and cDNA clones

One cotton cDNA clone designated pSKcF106A was isolated and sequenced (GenBank AF329635). The cDNA insert was found to be a partial cDNA clone of 791-bp derived from the 3′-half of the cognate *FAD2* mRNA, as determined by DNA sequence analysis by Dr. Irma Pirtle of our laboratory, and Dr. John Knesek of Texas Woman's University.

Initially, one *FAD2* genomic clone designated LCFg55 was isolated from a cotton (*Gossypium hirsutum*, cv. Acala SJ-5) genomic library harbored in the lambda vector EMBL3. Plaque hybridization was performed by the procedure of Benton and

Davis (1977). The bacterial strain used for screening the *FAD2* genomic DNA harbored in EMBL3 was *E. coli* K802. Approximately $1x10^7$ plaques of the cotton genomic library harbored in lambda EMBL3 was screened on 36 plates at a plaque density of $3x10^5$ plaque-forming units. Prehybridization was done at 55°C for 4 h in a solution containing 6xSSC, 0.5% SDS, 5x Denhardt's reagent, and denatured sheared salmon sperm DNA (100 μg/ml). Hybridization was done overnight in a solution of 20 mM Tris-HCl (pH 8.0), 2 mM EDTA, 2.5 mM Na pyrophosphate (pH 8.0), and a ^{32}P-labeled DNA fragment generated by random priming using a homologous 0.8-kb *Xho*I-*Eco*RI fragment from the cotton cDNA clone designated pSKcF106A as template DNA. The DNAs from intensely hybridizing positive plaques were initially purified by a mini-lysate procedure (Sambrook et al., 1989). The lambda clones containing possible *FAD2* genes were then digested with restriction endonucleases and analyzed by agarose gel electrophoresis and alkaline blot hybridization. The putative *FAD2* genomic clones with the most intense positive signals were selected and large phage DNA preparations done by the procedure of Tiemeier et al. (1977). Based on the tentative physical map, a 10-kb *Hin*dIII fragment of a genomic clone designated LCFg55 containing a putative *FAD2* gene was subcloned into the *Hin*dIII site of the cosmid vector pDELTA2 (Gibco BRL) for DNA sequence analysis. Since the clone LCFg55 was found to not contain a full-length *FAD2* gene, it became necessary to isolate an overlapping genomic clone.

A second cotton (*Gossypium hirsutum*, cv. Acala SJ-2) genomic library harbored in the Lambda FIXII vector (Stratagene), generously donated by Dr. Thea Wilkins (Univ. of California at Davis), was then screened to isolate additional clones

encompassing the entire *FAD2* gene. A unique, *FAD2* gene-specific 5′-flanking region fragment was generated by PCR amplification of 396 bp of the partial 5′-flanking region of the *FAD2* gene in LCFg55, and inserted in the vector pGEM7Zf(+) to prepare ^{32}P-labeled hybridization probe by random priming. A forward primer (5′-CCG AGC TCA CTA ATC TTC AAG TGT ATCC- 3′) and a reverse primer (5′-CGT CTA GAT GGT TAA AGT ACT CGG-3′) were designed based on the sequence of 568-bp of 5′-untranslated region of the *FAD2* gene from LCFg55. A 396-bp segment of the 5′-flanking region, beginning 37 nt upstream from the ATG initiation codon was amplified by PCR by Mr. Mongkol Nampaisansuk of this laboratory with a denaturation step of 94˚C for 30 sec, 30 sec at 53˚C for annealing, and 30 sec at 72˚C for primer extension for 30 cycles. The 396-bp segment of the 5′-flanking region was subcloned into the *Xba*I and *Sst*I sites of the vector pGEM7Zf(+) and designated FAD2pGEM. The *FAD2* 5′-flanking region insert in FAD2pGEM was sequenced by Dr. Irma Pirtle. The 396-bp insert of the recombinant plasmid was isolated by digestion of the plasmid DNA with *Xba*I and *Sst*I and purified by agarose gel electrophoresis and the QIAGEN QIAquick Gel Extraction Kit. This *Xba*I-*Sst*I fragment was used as template to generate ^{32}P-labeled probe by random priming to screen the Lambda FIXII cotton genomic library. Approximately $1x10^6$ plaques from the cotton genomic DNA library was screened by the plaque hybridization method of Benton and Davis (1977). The plaque lifts were hybridized at 60˚C overnight in a solution containing 6xSSC, 0.5% SDS, 10xDenhardt's reagent, and denatured sheared salmon sperm DNA (100 µg/ml) for 4 h. Hybridization was done overnight in a solution contaning 6xSSC, 0.5% SDS,

5xDenhardt's reagent, 20 mM Tris-HCl (pH 8.0), 2 mM Na_2EDTA (pH 7.5), 2.5 mM Na pyrophosphate (pH 8.0), denatured sheared salmon sperm (100 μg/ml), and the ^{32}P-labeled probe. Because of the identity with overlapping lambda genomic clone LCFg55, the plaques were screened using a more stringent condition (i.e. 60˚C instead of 55˚C). As a result, eight positive plaques were isolated and purified through two rounds of screening as mentioned above. The eight clones harbored in lambda FIXII were isolated using a minilysate phage preparation, and digested with *Bam*HI and *Sst*I (polylinker sites in the lambda FIXII arms).

IV. Large scale phage preparation

Two liter cultures of the lambda-sensitive bacterial strain *E. coli* K802 were inoculated with bacteriophage encompassing putative *FAD2* genes at a m.o.i. (multiplicity of infection) of 0.2 and grown overnight at 37°C. The lysate contained 1-2x10^{13} pfu/ml (plaque-forming units/ml) in a volume of 500 ml in four separate 2 L flasks. The lysed cultures were processed by adding 8 ml of chloroform with incubation at 37°C for about 10 min. The lysates were treated with pancreatic RNase A (1 μg/ml) and *E. coli* DNase I (1 μg/ml) and incubated at room temperature for 30 min. NaCl (29.22 g) was added to each 500 ml lysate to make the final concentration 1M, and the culture was placed on ice at 0°C for 2 h. Then the cellular debris and chloroform were removed by centrifugation at 5,150xg for 30 min. The phage was precipitated by adding polyethylene glycol (10% PEG 8000, w/v, Sigma), and placed overnight at 4˚C. The phage was centrifuged at 5,150xg for 30 min. The pellet was harvested and dissolved in 4 ml of SM buffer (20 mM Tris-HCl, pH 7.5, 20 mM

$MgSO_4 \cdot 7H_2O$, 100 mM NaCl , 2% gelatin). The phage particles were purified by CsCl density gradient ultracentrifugation (Sambrook et. al., 1989) as outlined below. The aqueous phase (7.5 ml) in SM buffer was transferred to 12 ml Oak Ridge tubes (Nalgene #3430-1610). The difference in weight of the empty and filled tubes were determined, and multiplied by a factor of 0.71 to determine the weight in grams of CsCl to add to each tube. The samples were centrifuged in a Beckman LS-65 ultracentrifuge at 38,000 rpm (22,000xg) for 18-24 h. After centrifugation, the white phage bands were collected with a Pasteur pipette. The phage bands were transferred to a dialysis bag (Fisher #08-667B with a 12,000 MW cut-off). The phage were dialyzed in 1 L of cold dialysis buffer (100 mM Tris-HCl (pH 8.0), 0.3 M NaCl) with gentle stirring overnight. The dialysis buffer was changed three times in order to completely remove CsCl from the phage preparation. The dialyzed phage were transferred into polypropylene centrifuge tubes and extracted with an equal volume of phenol/chloroform /isoamyl alcohol (25/24/1, v/v/v). The samples were centrifuged in the Sorvall SS-34 rotor at 5,000 rpm (2,790xg) for 15 min at 22°C. The top aqueous phase was transferred to a clean polypropylene tube and extracted with an equal volume of chloroform/isoamyl alcohol (24/1). After the aqueous phase of the chloroform extraction was removed, 2 volumes of cold 95% ethanol was mixed with the aqueous phase and placed at least two hours at -90°C in order to precipitate the phage DNA. The phage DNA was harvested by centrifuging at 12,500 rpm (14,400xg) for 20 min. Finally, the DNA pellet was dried and dissolved in a minimum volume of about 4 ml of TE buffer (1 mM EDTA, 1 mM Tris-HCl (pH 7.5)).

V. Physical mapping of phage containing putative cotton *FAD2* genes

The cotton genomic DNA inserts encompassing the potential *FAD2* genes in the lambda EMBL3 and Lambda FIXII vectors were cleaved with a variety of restriction endonucleases using either single or double digestions as described by Sambrook et al. (1989). The DNAs (about 1-4 μg) were digested with 10-20 units of restriction endonucleases in a final volume of 20 μl with the appropriate 10X buffer from the manufacturer at the appropriate temperature, usually 37ºC. Complete digestion took 3-24 h depending on the activity of the particular restriction endonuclease. The restriction enzyme activity, compatible buffers, and inactivation time were specified by the manufacturers. DNA precipitation with 95% ethanol may have been required prior to the second digestion if a compatible buffer could not be found. After complete digestion, a portion of DNA from each reaction was separated and sized by agarose gel electrophoresis. The protocol involved loading the digested DNA into 0.8-2.0 % agarose gels in TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH 8.5). Next, the gel was placed in a horizontal gel electrophoresis apparatus and submerged in TAE buffer. The digested DNA was mixed with loading dye (containing 0.25% bromphenol blue, 0.25% xylene cyanol, 1% sodium dodecyl sulfate, 0.1 M Na_2EDTA, pH 8, and 20% Ficoll 400). Gel electrophoresis was done at 34 volts overnight. The gel was stained with ethidium bromide (0.5 μg/ml) and photographed when exposed to UV light. The sizes of all DNA fragments resulting from single and double digestions with a variety of restriction endonucleases were determined, and the physical maps of the *FAD2* clones harbored in the lambda vectors EMBL3 and Lambda FIXII were constructed. For

determining the location of the *FAD2* coding region, a ^{32}P-labeled DNA fragment generated by random priming using the 0.8-kb *Eco*RI/*Xho*I fragment of cotton *FAD2* cDNA from the clone pSKcF106A as template DNA was hybridized to alkaline blots of the DNA which had been digested with a variety of restriction endonucleases.

VI. Subcloning of the cotton *FAD2* genomic DNA

Appropriately-sized DNA fragments containing the flanking regions and the coding region from the *FAD2* genomic clones were subcloned into the vectors pDELTA2 or pUC19, depending on the purpose of the study. For deletion subclones, an appropriate DNA fragment from the lambda genomic clone LCFg55 was subcloned into the vector pDELTA2 for automated sequencing by Dr. John Knesek of Texas Woman's University, Denton, TX and designated pCFg55. Also, a *Hin*dIII and a *Sal*I DNA fragment from the *FAD2* genomic clone LCFg24 in λ FIXII were subcloned into pUC19 for sequencing and designated pCFg24H and pCFg24S, respectively.

For pDELTA2 subcloning, about 200-300 ng of DNA from the putative cotton *FAD2* gene harbored in the genomic clone designated LCFg55 was digested with *Hin*dIII and purified by electrophoresis on a 0.8% LE agarose gel. A 10-kb *Hin*dIII fragment was cut from the agarose gel. The DNA fragment was then purified by a QIAGEN Gel Extraction Kit protocol (QIAGEN, Valencia, CA). The DNA was redissolved in H_2O and the DNA concentration determined by electrophoresis on a 0.8% LE agarose gel. The vector (pDELTA2) was also digested with *Hin*dIII, ethanol precipitated, and its concentration estimated by gel electrophoresis on a 0.8% LE agarose gel. The *Hin*dIII fragment of pDELTA2 was dephosphorylated using calf

intestine alkaline phosphatase (CIP) (GIBCO BRL). The reaction mixture (20 µl of pDELTA2 DNA (15 fmol), 2 µl of 10xCIP buffer and 2 units of CIP) was incubated at 37°C for 60 min. The *Hin*dIII-digested vector was phenol-extracted, ethanol-precipitated, and redissolved in H_2O for ligation to the *Hin*dIII fragment from LCFg55. The ratio of insert to vector DNA used for ligation was 3 to 1. The ligation mixture consisted of 45 fmole of *Hin*d III-digested lambda DNA fragments, 15 fmole of pDELTA2 DNA, 4 µl of 5x ligase buffer, 1.0 µl (0.1 units) of T_4 DNA ligase (GIBCO BRL), and H_2O in a final volume of 20 µl. The ligation mixture was incubated at 23-26°C for 1h. The reaction was terminated by freezing at –20°C.

For subcloning a 6.8-kb *Hin*dIII DNA fragment from the *FAD2* genomic DNA clone LCFg24 harbored in λFIXII, LCFg24 DNA was digested and the 6.8-kb fragment purified by 0.8% gel electrophoresis and purified with the QIAGEN Gel Extraction Kit. The purified DNA was precipitated with ethanol, the same as in the purification method used for the fragment in pDELTA2. Also, the pUC19 vector DNA was digested with *Hin*dIII, dephosphorylated using CIP, phenol extracted, and ethanol precipitated. The 6.8-kb insert containing the putative *FAD2* gene and the *Hin*dIII-digested pUC19 DNA were ligated using a ratio of 3 to 1 (insert to DNA) and T_4 DNA ligase. The mixture consisted of the following: 300 ng of insert (4 µl), 100 ng of vector (5 µl), 10 µl of ligase buffer (50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% (w/v) polyethylene glycol-8000), and 0.1 units of T_4 DNA ligase (1 µl). The mixture was incubated at 25°C for 1 h. In addition, for subcloning of a fragment with the 5´-untranslated region of LCFg24, a 4.1-kb *Sal*I fragment and the *Sal*I digested-pUC19

DNA were ligated by Mr. Mongkol Nampaisansuk and Dr. Robert M. Pirtle of this laboratory, using the same conditions as described above.

The transformation of competent *E. coli* cells was done by electroporation. Electrocompetent *E. coli* DH10B and DH5α cells were prepared by culturing in LB medium (Luria-Bertani medium: 10 g Bactotryptone, 5 g yeast extract, and 5 g NaCl per liter) at 37°C until the optical density at 600 nm was between 0.5 and 1.0. The cells were centrifuged at 5,500 rpm (2,790xg in the SS-34 roter) at 4°C for 15 min. Then the cells were resuspended in one volume of sterile cold distilled water. The cells were then centrifuged twice at 5,500 rpm (2,790xg in the SS-34 rotor) at 4°C and resuspended in 0.5 volume of sterile cold H_2O the first time and 0.02 volume of sterile cold H_2O the second time. Lastly, the pellets were resuspended in 0.002-0.003 volume of sterile cold 10% filtered glycerol. The electrocompetent *E. coli* DH10B and DH5α cells (0.2 ml) were distributed into microfuge tubes and frozen at -90°C. Before use, the *E. coli* DH10B and DH5α cells were thawed on ice for 5-10 min.

The transformation mixtures were prepared by mixing 100 ng of plasmid DNA in 1-2 µl H_2O (diluted 1:5 from the ligation mixture) and 35 µl of the electrocompetent *E. coli* DH10B or DH5α cells. The transformation mixtures were transferred to BTX disposable cuvettes with a 1 mm gap. The suspension was checked to make sure it touched both side walls of the cuvette. The cuvettes were chilled on ice for 1 min prior to electroporation. The samples were then electroporated at 1.5 kV with an electric field strength of 13.0-15.0 kV/amp for 5-6 millisecond. After electroporation, 960 µl of LB broth was immediately added into the samples, which were mixed thoroughly using

a pipette. The samples (50-100 µl) were then transferred to polypropylene tubes and shaken at 225 rpm for 1 h and then spread on selection plates. The transformed *E. coli* DH10B cells were spread on LB agar plates containing ampicillin, X-gal (5-bromo-4-chloro-3-indoyl-β-D galactoside) and IPTG (Isopropylthio-β-D-galactoside). The plates were incubated at 37°C for 24 h, and the white colonies were selected using the blue-white system of selection of the interrupted *lacZ* gene in the pDELTA2 and pUC19 vectors. After incubation for two days, the white colonies of the pDELTA2 subclones in *E. coli* DH10B cells were streaked onto fresh LB agar plates containing ampicillin, kanamycin, and X-gal in order to confirm the recombinant clones. The inserts of the plasmids containing the potential *FAD2* coding region and 5′-flanking region were tested by digestion with appropriate restriction enzymes and agarose gel electrophoresis after rapid plasmid DNA preparations.

Rapid plasmid preparations

Potential pDELTA2 transformants containing the *FAD2* gene were grown on the ampicillin/kanamycin/chloramphenicol plates for two days (to ensure that the colonies would survive on LB medium containing the antibiotics ampicillin, kanamycin, and chloramphenicol in order to isolate potential clones with *Tn*1000 transposons (Gibco BRL)). In contrast, in screening the pUC19 transformants potentially containing a larger 5′-flanking region of the *FAD2* gene, the transformants were isolated in LB medium containing Amp/X-gal/IPTG. The selectable-single colonies for pDELTA2 and pUC19 recombinant plasmids having the *FAD2* gene and 5′-flanking region were grown in 5 ml of LB/Amp/Kan and LB containing ampicillin, respectively, overnight at 37°C.

The cell cultures (1.5 ml) were pelleted at 10,000 rpm (4,510xg) for 1 min in a microcentrifuge, and the supernatants were removed by Pasteur pipette. The pellets were resuspended in 100 µl of ice-cold lysis buffer (0.9% glucose, 10 mM Na_2EDTA, 25 mM Tris-HCl, pH 8.0). Denaturing solution (200 µl of 0.2 N NaOH, 1% sodium dodecyl sulfate) was added to each mixture, and then thoroughly mixed by inversion. Then all mixtures were placed on ice for 5 min. The lysates were mixed with 150 µl of ice-cold potassium acetate, pH 4.8 (3 M potassium acetate, 11.5% glacial acetic acid) and gently mixed by inversion for 10 min. The lysates were placed on ice for 5 min. Then the cell debris and chromosomal DNA were pelleted in a microcentrifuge for 1 min. The supernatant was transferred to a fresh tube and treated with DNase-free RNase A to a final concentration of 20 µg/ml at 37°C for 20 min. In order to do a phenol/chloroform extraction, one volume of TE-saturated phenol/chloroform (1:1) was added to the supernatants, and the samples were vortexed for 1 min and centrifuged at 10,000 rpm (4,510xg) for 5 min. The upper aqueous phase was transferred to a fresh tube and one volume of chloroform/isoamyl alcohol (24:1) added. The samples were vortexed for 1 min and the DNA precipitated by the addition of 2.5 volumes of 100% ethanol, and letting stand at -20°C for 5 min. The plasmid DNA was pelleted by centrifugation at 10,000 rpm (4,510xg) for 5 min, and rinsed with 70% ethanol. Then the pellets were dried under vacuum. Each dried pellet was dissolved in 10 µl of sterile deionized water. The expected yield of plasmid DNA was 1-4 µg per 1.5 ml of culture. The samples were characterized by digestion with restriction endonucleases and electrophoresis on 0.8% agarose gels.

Large scale plasmid preparation

After it was verified that the recombinant clones from the subcloning step had the *FAD2* gene insert with the appropriately-sized 5′-flanking region, the recombinant plasmids in both *E. coli* DH10B and DH5α cells were extracted by an alkaline lysis method. LB broth (2-3 liters) containing ampicillin and kanamycin or containing only ampicillin was prepared for extraction of recombinant plasmids. Cultures of 6 ml of LB broth containing the appropriate antibiotics (as described above) were grown overnight at 37°C until log phase. The entire 6 ml culture was used to inoculate a 500 ml culture in LB broth with the appropriate antibiotics in 2 liter flasks. The large 500 ml cultures were grown at 37°C until the optical density at 600 nm was equal to 0.4. The cells were harvested by centrifugation for 10 min at 4°C at 4,000 rpm (1,480xg) in a Sorvall GS-3 rotor. The bacterial pellets were weighed before and after being washed in 100 ml of STET (8% sucrose, 0.5% Triton X-100, 50 mM EDTA, 5 mM Tris-HCl, pH 8.0) in order to determine the volume of each of the following solutions that should be added. The pellets (4-8 gram of cells per tube) were resuspended in 22.5 ml of glucose/Tris-EDTA solution (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl (pH 8.0)), and 2.25 ml of freshly prepared lyzozyme (10 mg/ml in 10 mM Tris-HCl (pH 8.0)) was added to the cell suspensions. Freshly prepared NaOH/SDS solution (0.2 N NaOH, 1% sodium dodecyl sulfate) was added and the solution mixed by gentle inversion. The mixture was neutralized by the addition of 22.5 ml of 5 M potassium acetate, pH 4.8 (3 M potassium acetate, 11.5% glacial acetic acid), and allowed to stand on ice for 10 min. The cellular debris was pelleted at 4,000 rpm (1,480xg) for 30 min at 4°C. The

supernatants were collected and precipitated by addition of 0.6 volumes of isopropanol at room temperature for 10 min. The precipitated DNAs were pelleted by centrifugation in the Sorvall GS-3 rotor at 6,000 rpm (3,320xg) for 15 min at room temperature. The pellets were resuspended in 10 ml of TE buffer. RNase A was added to the DNA solution to a final concentration of 10 μg/ml, and the solution was incubated at room temperature for 1 h. A phenol-chloroform extraction was performed by the addition of one volume of phenol/chloroform/isoamyl alcohol (25:24:1), with vigorous vortexing. Then the samples were centrifuged in the Sorvall SS-34 rotor at 6,000 rpm (3,320xg) for 15 min. The aqueous phase was transferred to a new tube and one volume of chloroform/isopropanol alcohol (24:1) was added to the aqueous phase. The samples were vigorously vortexed for 2 min, and centrifuged in the SS-34 rotor at 6,000 rpm (3,320xg) at room temperature for 15 min. The aqueous phases were transferred to clean tubes. A 0.1 volume of 3 M sodium acetate, (pH 5.2) and 2.0 volumes of 100% ethanol were mixed with the aqueous phase for precipitation overnight at –90°C. The ethanol-precipitated samples were warmed to 4°C on ice and then centrifuged in the Sorvall SS-34 rotor at 4°C and 12,500 rpm (14,400xg) for 30 min. The pellets were washed with 10 ml of cold 70% ethanol, dissolved in an appropriate volume of 1xHPLC starting buffer (25 mM Tris-HCl, 1 mM EDTA, pH 8.0), and then purified by Dr. Irma Pirtle by HPLC using the general conditions of Merion and Warren (1989). The absorbances of the crude plasmid DNA solution was measured at 260 nm and 280 nm to estimate the amount and purity of the plasmid DNA sample.

VII. Generating Nested Deletions

The pCFg55 subclone containing a *FAD2* gene was used to create deletions by recombination using a genetically engineered transposon system (*Tn*1000) available from Gibco BRL Life Sciences (Wang et al., 1993; York et al., 1998). The deletion end points were mapped by determining the sizes of numerous subclones by agarose gel electrophoresis. At the endpoint of any deletion by the transposon, the primer binding sites served as unique starting reference points for sequencing using fluorescent-labeled primers. The DNA fragments were analyzed on a LICOR 4000L semi-automated DNA sequencer at Texas Woman's University, Denton, TX, by Dr. John Knesek.

This nested-deletions method is an approach for sequencing large DNAs without using numerous synthetic oligonucleotide primers. The approach is to bring different sections of target DNA to a constant site after the transposon *Tn*1000 integrates randomly in the target DNA sample. Clockwise and counterclockwise subclones are created depending on the loss of selectable markers. The random transposition events lead to the nested deletions that extend in varying distances from one transposon end through the selectable marker, and into, but not beyond, the cloned insert DNA. The selectable markers for clockwise subclones are sucrose resistance (Suc^r) and ampicillin resistance (Amp^r). Streptomycin resistance (*strA*) and kanamycin resistance (Kan^r) are the selectable markers for counterclockwise subclones. Thus, the SP6 promoter universal primer for sequencing the clockwise subclones and the T7 promoter universal primer for sequencing the counterclockwise subclones (GIBCO/BRL, Deletion Factory System 2.0 manual) were used.

The recombinant plasmid from pDELTA2 designated pCFg55 (40 µg or 60 pM) was transformed into 100 µl of electrocompetent *E. coli* DF1 cells by electroporation at 1.5 kV. The transformation mixture (100 µl) was spread onto a LB ampicillin/chloramphenicol/kanamycin plate which was incubated overnight at 37°C. On the second day, three to five white colonies of the transformed cells were each added to 2 ml cultures in of LB/ampicillin broth. Another three to five white colonies were used to inoculate 2 ml cultures in of LB/kanamycin broth. The plasmid DNAs from both LB/ampicillin and LB/kanamycin cultures were purified using the mini-plasmid isolation protocol (Sambrook et al., 1989). The identity and quality of the plasmid DNA samples were analyzed by agarose gel electrophoresis. The concentration of pCFg55 plasmid DNA in each sample was visually estimated based on the intensities of the bands relative to the intensities of known amounts of standard DNA fragments and the amount of sample DNA loaded. Maximum efficiency *E. coli* DH10B competent cells (50 :l) were transformed with 2.5 µl (0.15 fmol) of the diluted DNA in LB/ampicillin broth. An aliquot (100 µl) of each transformation reaction was streaked onto a LB/sucrose/ampicillin deletion selection plate. Another 2.5 µl (0.15 fmol) of the diluted DNA from the LB/kanamycin culture was used to transform 50 µl of Maximum Efficiency *E. coli* DH10B (GIBCO/ BRL) competent cells. An aliquot (100 µl) of the transformation reaction was spread onto a LB streptomycin/kanamycin deletion selection plate. In order to prepare for DNA analysis using a mini-plasmid preparation procedure (Sambrook et. al., 1989), the deletion subclones from both sucrose/ampicillin and streptomycin/kanamycin selection plates were used to inoculate 2 ml cultures of

colonies from the LB/sucrose/ampicillin and LB/streptomycin/kanamycin selection plates for the clockwise and counterclockwise deletion subclones, respectively. The DNAs from deletion subclones of the clockwise and counterclockwise constructs were purified by the Wizard Plus Minipreps DNA Purification System (Promega) to prepare the DNAs for sequence analysis. Based on the protocol for transposon deletion subcloning (GIBCO/BRL, Deletion Factory System 2.0 manual)**,** two colonies are supposedly required per kilobase of insert DNA for a 0.5-kb deletion sequence interval. However, in practice, it was necessary for us to have 10-20 colonies per kilobase. Each deletion subclone was characterized by sizing on agarose gels. The plasmid DNAs of both clockwise and counterclockwise subclones were purified and the concentrations visually estimated by comparisons to known standard DNAs on agarose gel electrophoresis.

VIII. DNA sequencing

The nucleotide sequence of the clockwise and counterclockwise deletion plasmids for the *FAD2* genes in pCFg55 and the overlapping 6.8-kb *Hin*dIII and 4.1-kb *Sal*I fragments with the *FAD2* gene in pCFg24H and pCFg24S were determined by Drs. John Knesek and Irma Pirtle, using semiautomated or manual sequencing, respectively. The nucleotide sequences of the pCFg24H and pCFg24S DNA was done using a primer-based approach. For semiautomated sequencing, the SequiTherm EXCEL II Long-Read DNA Sequencing kit (Epicentre Technologies) was used with a LI-COR automated DNA sequencing machine and fluorescently-labeled primers. Fluorescently-labelled primers corresponding in sequence to those of the SP6 promoter and T7 promoter of

pDELTA2 were used as forward and reverse primers. Manual DNA sequencing was done by the chain termination method with a thermostable DNA polymerase (Thermo Sequenase from Amersham) and ^{33}P-labeled dideoxyribonucleotides for terminator cycle sequencing (Fan et al., 1996). The DNA sequences of each of the overlapping deletion subclones was analyzed by DNASIS software (Hitachi). The non-overlapping and GC-rich compressed regions of pCFg55 were sequenced using manual sequencing with the chain termination method (Thermo Sequenase from Amersham). The DNA sequences derived from semi-automated DNA sequencing of the deletion subclones and the manual DNA sequencing of the nonoverlapping and compressed areas were assimilated and analyzed for closure by Dr. Irma Pirtle of this laboratory using DNASIS software (Hitachi).

IX. Genomic DNA gel blot analysis

The copy number or reiteration frequency of *FAD2* genes in the cotton genome was determined by genomic blot hybridization. Cotton genomic DNA of *Gossypium hirsutum* was extracted from young cotton leaves according to the procedure of Paterson et al. (1993) by Mr. Jeffery Wilkinson of our laboratory. Cotton genomic DNA (15 µg per reaction) was digested with several restriction enzymes and electrophoresed through 0.8% agarose gels. The gels were soaked in depurination solution (0.25 M HCl) for 45 minutes, and then denatured in 0.5 M NaOH for 20 min. The gels were blotted onto positively-charged nylon membranes (Amersham Hybond N^+). Prehybridization of the nylon membranes was performed at 60°C in a solution containing 6xSSC, 0.5% SDS, 10X Denhardt's reagent, 0.2% Ficoll (Type 400, Pharmacia), 0.2%

polyvinylpyrrolidone, 0.2% bovine serum albumin (Fraction V; Sigma) and sheared salmon sperm DNA (100 μg/ml) for 4 h. Hybridization was done using ^{32}P-labeled DNA fragments generated from random priming of a homologous cotton *FAD2* cDNA fragment (the 0.8 kb *Eco*RI/*Xho*I fragment from pSKcF106A) under high-stringency conditions. Hybridization was done in a solution containing 6xSSC, 0.5%SDS, 5x Denhardt's reagent, 20 mM Tris-HCl (pH 8.0), and sheared salmon sperm DNA (100 μg/ml) overnight at 60ºC. Then the filters were washed with a series of buffers containing (1) 2xSSC; (2) 2xSSC, 0.1%SDS; (3) 1xSSC, 0.1%SDS; and (4) 0.5xSSC, 0.1%SDS at 65°C for 30 min per buffer.

X. Expression of the cotton *FAD2* gene in yeast cells

The open reading frame of the cotton *FAD2* gene was used as a template by Mr. Mongkol Nampaisansuk of this laboratory for PCR amplification to subclone the *FAD2* gene into the yeast-bacterial shuttle vector pYES2 (Invitrogen). The forward amplimer, 5′-GGG AGC TCA TGG GTG CAG GTG GCA GAA-3′ was used to create a *Sst*I site (underlined) adjacent to the *FAD2* ATG initiation codon in the 5′-flanking region of the open reading frame. The reverse amplimer, 5′-GCG AAT TC T TAG ATC TTA TTT CTA AAC CAA AAT ACA CC-3′ was designed to provide a *Eco*RI (underlined) site in the 3′-flanking region of the open reading frame. The 1.2-kb PCR product was digested with *Eco*RI and *Sac*I, and inserted into the *Eco*RI/*Sac*I-digested pYES2 to generate the recombinant DNA plasmid pYES/FAD2. The plasmid pYES2 is a *Saccharomyces cerevisiae* expression vector which contains the *GAL1* promoter, and is useful for inducible expression of genes inserted into the polylinker multiple cloning

site of this vector. The pYES/FAD2 plasmid DNA was sequenced by Dr. Irma Pirtle of this laboratory to confirm that the coding region was identical to that of the *FAD2* gene and that the construct was in frame relative to the *GAL1* promoter. The entire coding region of *FAD2* was found to be in proper sense orientation relative to the *GAL1* promoter. The *S. cerevisiae* strain INVSc1 (Invitrogen, MATa/a, *his 3…1, leu2/leu2, trp1-28a/trp1-28a, ura 3-52*) was transformed with both pYES2 DNA and pYES2/FAD2 DNA by electroporation and the lithium acetate method (Gietz et al., 1992). Recombinant yeast cells were selected on uracil-deficient agar plates in 0.67% yeast nitrogen base without amino acid, 2% glucose, 0.01% each of adenine, arginine, cysteine, methionine, phenylalanine, proline, serine, tyrosine, and valine, all obtained from Sigma Chemical Co. (St. Louis, MO).

Lithium transformation

Competent yeast cells (strain INVSc1 from Invitrogen) were prepared by culturing a pure colony in 10 ml of Yeast Extract Peptone Dextrose (Sigma) medium (YPD, containing 1% yeast extract, 2% peptone, 2% dextrose), and incubating overnight at 30°C. The cells were diluted to an OD_{600} of 0.4 in 50 ml of YPD medium and grown for an additional 2-4 h. The cells were pelleted at 5,500 rpm (2790xg) and resuspended in 40 ml 1xTE buffer (10 mM Tris, 1 mM EDTA, pH 7.5). The cells were centrifuged at 5,500 rpm (2,790xg) and resuspended in 2 ml of 1xLiAC/0.5xTE (100 mM lithium acetate, pH 7.5, 5 mM Tris-HCl, pH 7.5, 0.5 mM EDTA). The resuspended yeast cells were incubated at room temperature for 10 min. The lithium transformation reaction contained the following: 1 μg of plasmid DNA, 100 μg of

denatured sheared salmon sperm DNA, and 100 µl of the yeast cell culture. The cells were incubated at room temperature for 10 min. Then, 700 µl of 1xLiAc/40%PEG-3350/1xTE (100 mM lithium acetate, pH 7.5, 40% PEG-3350, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA) was added to the transformation reaction mixture and incubated at 30°C for an additional 30 min. Then, 88 µl of DMSO (dimethyl sulfoxide) were added and the yeast cells were subjected to heat shock at 42°C for 7 min. The transformation mixture was centrifuged for 2 min and the supernatant was removed. The yeast cell pellet was resuspended in 1 ml of 1xTE and re-pelleted. The 50-100 µl of cell suspension in 1xTE was plated on a selection medium called SC-U medium (synthetic complete minus uracil) consisting of 0.67% yeast nitrogen base without amino acids, 2% carbon source (raffinose or glucose), 0.01% each of adenine, arginine, cysteine, leucine, lysine, threonine, and tryptophan, and 0.005% each of aspartic acid, histidine, isoleucine, methionine, phenylalanine, proline, serine, tyrosine, and valine.

Electroporation of Yeast cells

Electrocompetent yeast cells were prepared as described in Becker (1991) and in the Genetronics, BTX instruction manual. *Saccharomyces cerevisaie* strain INVSc1 was grown in 500 ml of YPD (1% yeast extract, 2% Bacto peptone, 2% glucose) in a 2 L flask at 30°C with vigorous shaking. When the cell density reached an OD_{600} of 1.3-1.5 (about $1x10^8$ cells/ml), the yeast culture was centrifuged at 5,500 rpm (2,790xg) for 5 min and resuspended in 100 ml YPD broth with 2.0 ml of 1 M HEPES-NaOH (pH 8.0). Sterile 1 M dithiothreitol (2.5 ml) was added while swirling gently. The yeast cell suspension was incubated at 30°C for 15 min. The suspended yeast cells were added to

400 ml of sterile ice-cold distilled H_2O to bring the volume to 500 ml. The yeast suspension was pelleted at 5,500 rpm (2,790xg) at 4°C for 5 min, and resuspended in 500 ml of sterile ice-cold ddH_2O. After the first washing with H_2O, the cells were spun at 5,500 rpm (2,790xg) at 4°C for 5 min, and resuspended in 250 ml of sterile ice-cold H_2O. The third wash was spun at 5,500 rpm (2,790xg) at 4°C for 5 min, and resuspended in 20 ml of sterile cold 1 M sorbitol. Then the yeast cells were pelleted by centrifugation at 4°C for 5 min. Finally, the electrocompetent yeast cells were resuspended in 0.5 ml of sterile cold 1 M sorbitol, and 100 µl aliquots were pipetted into 1.5 ml microfuge tubes, and stored at –90°C.

The competent yeast cells were thawed on ice for 10 min, and distributed in 40 µl aliquots for transformation. Either 50 ng of pYES2 or pYES2/FAD2 plasmid DNA (no less than 100 ng) in a maximum of 5 µl was added into a 1.5 microfuge tube. The yeast electrocompetent cells and either pYES2 or pYES2/FAD2 plasmid DNAs were mixed gently, and incubated on ice for about 5 min. Then the cell mixtures were transferred to a cold 2-mm sterile electroporation cuvette (BTX disposable cuvette P/N 620), and tapped to the bottom of the cuvette. The samples were pulsed at 1.5 kv 0.25 µF 129Ω (estimated field strength about 7.5 kv/cm). After pulsing, 1 ml of cold 1M sorbitol was immediately added, and the yeast cells were transferred with gentle mixing to a cuvette.

Yeast induction by galactose

A single colony of the yeast strain INVSc1 harboring either pYES2 or pYES2/FAD2 plasmid was grown at 30°C overnight in 25 ml of SC-U medium, pH 5.8

(0.67% yeast nitrogen base without amino acids, 2% carbon source (raffinose or glucose), 0.01% each of adenine, arginine, cysteine, leucine, lysine, threonine, and tryptophan, 0.005% each of aspartate, histidine, isoleucine, methionine, phenylalanine, proline, serine, tyrosine, and valine). The cells were pelleted by centrifugation at 5,500 rpm (2,790xg) for 5 min at 42°C. The amount of overnight culture (with approximately an OD_{550} of 0.8) in 25 ml of SC-U medium containing 2% glucose was transferred to a 50 ml polypropylene tube. The yeast cells were pelleted at 5,500 rpm (2,790xg) for 10 min at 4ºC, washed twice with TE buffer, and centrifuged at 4,500 rpm for 10 min at 4ºC. The cells were resuspended in 1 ml of induction medium (SC-U medium containing 2% raffinose and 2% galactose) with or without oleic acid. Then, the cells were grown at 30ºC with shaking for three generations. For induction with oleic acid, the cell resuspension was added to 25 ml of SC-U medium containing 2% raffinose, 2% galactose, 1% tergitol type NP-40 with oleic acid as substrate to a final concentration of 0.5 mM (Aki et al., 1999). The cell cultures were incubated at 30ºC with shaking for three generations until the stationary phase was reached. It has been shown (Kajiwara et al., 1996) that the required incubation time for the *FAD2* transcript to be detected was at least 21 h (three generations). During harvesting, 1 ml of the yeast culture was removed from each flask and its OD_{550} determined. Next, the yeast cells were washed three times with 1 ml of ice-cold distilled H_2O in order to remove oleic acid and metabolites that could interfere with the next process. The yeast pellets were resuspended in 800 µl of ice-cold distilled H_2O, and stored at –90ºC until ready for use. Then, fatty acid components in the yeast cell were extracted, and transmethylated by Dr.

Kent Chapman of our Department. The fatty acid methyl esters were analyzed by gas-liquid chromatography (GC) as described by Chapman and Trelease (1991).

CHAPTER III

RESULTS

Gossypium hirsutum is a strain of cotton which is grown for its seed hairs since it is used for textile fiber. Also, other cotton products, especially cotton seed, are commercially useful in food and nonfood markets. Thus, success in designing new cotton oils can provide higher-value products of cotton seed for the seed oil industry. Since most fatty acid desaturases (FAD2, FAD3, FAD6, FAD7) are integral membrane proteins, protein characterization and purification by normal biochemical purification approaches is very difficult. Also, there is evidence that the FAD2 desaturase is likely located in the endoplasmic reticulum. For example, the mutants of *Arabidopsis* at the *FAD2* locus have a reduced activity of the endoplasmic reticulum desaturase (Miguel and Browse, 1992). Thus, cotton FAD2 likely exists in the endoplasmic reticulum, and is probably very difficult to purify using standard biochemical isolation procedures. Therefore, this project involves using the procedures of molecular biology and genetic engineering to clone and characterize the *FAD2* genes in the cotton genome.

Expression of fatty acid desaturase genes may be one of the control points in fatty acid metabolism that is required to modify the fatty acyl chain in storage lipids and membrane phospholipids (Ohlrogge, 1994). To better understand the mechanisms that regulate fatty acid desaturation in cotton, a fatty acid desaturase (*FAD2*) gene was isolated and characterized. In the future, we may introduce these fatty acid desaturase genes into transgenic cotton in order to improve the nutritional or economic value of

cotton oils and the cold tolerance of the cells in plants. The analysis of *cis*- and *trans*-acting promoter/enhancer regulatory elements for the desaturase genes will help us to understand how cotton adjusts its gene expression for cold tolerance as well as disease resistance. How desaturases are expressed during seed development may be controlled by these *cis*- and *trans*- regulatory elements. Thus, this research involved the isolation of a ω-6 desaturase gene (*FAD2*) by screening a cotton genomic DNA library. The gene structure was characterized by physical mapping, DNA sequencing, and genomic blot hybridization. Also, the expression of this *FAD2* gene was analyzed. All this information will assist us in analyzing and potentially modifying the *FAD2* gene in cotton, so that the modification of cotton seed oil fatty acid and modifications of the fatty acid compositions of membrane phospholipids will be possible in the future.

I. Screening of the cotton cDNA and genomic DNA libraries

Because of the functional divergence of plant ω-3 and ω-6 desaturases, the cotton FAD2 protein shares a high degree of amino acid sequence identity with other plant FAD2 proteins. For example, the alignment of the deduced amino acid sequences for plant desaturase-related sequences from rapeseed and soybean show a 72% identity (Hitz et al., 1994). In addition, sequence homology between the cDNA sequences of soybean *FAD2-1* and *FAD2-2* cDNAs have significant homology with the *Arabidopsis FAD2* gene (Okuley et al., 1994). Thus, there are 75%-85% identities among the plant desaturases, and the cotton *FAD2* gene may well have a high identity (75%) to similar genes in other plants.

The cotton cDNA library generated from cotyledon mRNA of 48-hour dark-grown cotton seedlings was screened with DNA fragments generated by random priming from an *Arabidopsis* cDNA clone designated *FAD2-43* (GenBank # T13887, clone # 43A5T7) which was obtained from Dr. R.N. Trelease of Arizona State University (Ni and Trelease, 1991). A 0.5-kb *Sal*I fragment of the plasmid DNA *FAD2-43* (GenBank # T13887) was used as template to generate random-primed probe for screening several hundred thousand plaques from the cotton cDNA library by the plaque hybridization method of Benton and Davis (1977). A plasmid DNA containing a potential *FAD2* cDNA named pSKcF106A was isolated and purified by HPLC. The restriction digestion and agarose gel electrophoresis of the plasmid DNA revealed that this cotton cDNA was a partial-length cDNA clone, since it only contained a 0.8 kb *Xho*I/*Eco*RI insert of a cotton *FAD2* cDNA. The cDNA sequence was determined by Dr. John Knesek at Texas Woman's University and is deposited in GenBank (Accession Number AF329635). The 0.8-kb *Xho*I/*Eco*RI fragment was used as template to prepare ^{32}P-labeled probe by random priming by Dr. Irma Pirtle of our laboratory. This homologous cotton probe was used to screen a cotton genomic DNA library and physically map the *FAD2* genomic lambda DNA designed LCFg55. The 0.8-kb *XhoI*/*Eco*RI fragment excised from the cotton *FAD2* cDNA clone was also used to prepare ^{32}P-labeled DNA probe to screen a cotton (*G. hirsutum*, cv. Acala SJ-2) genomic DNA library provided by Dr. Thea Wilkins of the University of California at Davis.

A cotton (*G. hirsutum*, cv. Acala SJ-5) genomic DNA library harbored in the lambda vector EMBL3 was screened by hybridization with the ^{32}P-labeled probes, either the 0.8-kb *Xho*I/*Eco*RI fragment of the cotton *FAD2* cDNA or the *Arabidopsis* cDNA fragment (*FAD2-43*). The prehybridization and hybridization was done at 55˚C using the hybridization conditions described in Methods and Materials. Twelve positive plaques were isolated and purified by a minilysate phage preparation procedure. The 12 *FAD2* genomic clones were digested with *Bam*HI, *Eco*RI, and *Sal*I, which were restriction sites in the polylinker region of lambda EMBL3, as shown in the agarose gel shown in Figure 5. Alkaline blot hybridization of these 12 genomic DNA fragments was done using the ^{32}P-labeled 0.8-kb *Xho*I/*Eco*RI fragment of the cotton cDNA clone pSKcF106A as probe. Ten to twelve of the cotton genomic clones gave an intense positive hybridization signal after autoradiography (Figure 6). However, only three of the genomic clones were chosen to be the representatives for the different forms of the *FAD2* genomic clones. The three clones were designated LCFg55A.aa, LCFg62B.aa and LCFg63A.bb. These three cotton genomic clones were purified using a large scale phage isolation procedure. The optimum conditions used for amplification was at an m.o.i. of 0.2, or the ratio of the phage to bacteria was 1:5. Also, a titer of the master phage stock of approximately 10^{13} to 10^{14} pfu/ml (plaque-forming units) was used in order to obtain complete lysis for the large scale phage preparation. The optimal temperature for the growth of the bacteriophage was 39˚C. As a result, the yield of the phage containing putative *FAD2* clones could be maximized to 10^{10} to 10^{11} pfu/ml prior to the purification step.

Figure 5. Agarose gel electrophoresis of putative *FAD2* genomic clones harbored in the phage EMBL3. DNAs from the 12 genomic clones (about 1.5 μg of DNA per reaction) were digested with *Bam*HI, fractionated on a 0.8% agarose gel, and transferred to a positively- charged nylon membrane. Initially, the clones had an arbitrary designation of FE (shown in this figure) which was changed to LCF in the later work. The clones gave five different profiles of DNA fragments. The clones designated LCF55A.aa, LCF55A.ab, LCF55A.ac, LCF62A.ca, LCF63A.aa, and LCF63A.ba were judged to be the same clones. The clones designated LCF62B.aa, LCF63A.bb, LCF64A.aa, and LCF64A.ba were found to be different clones from their DNA digestion profiles.

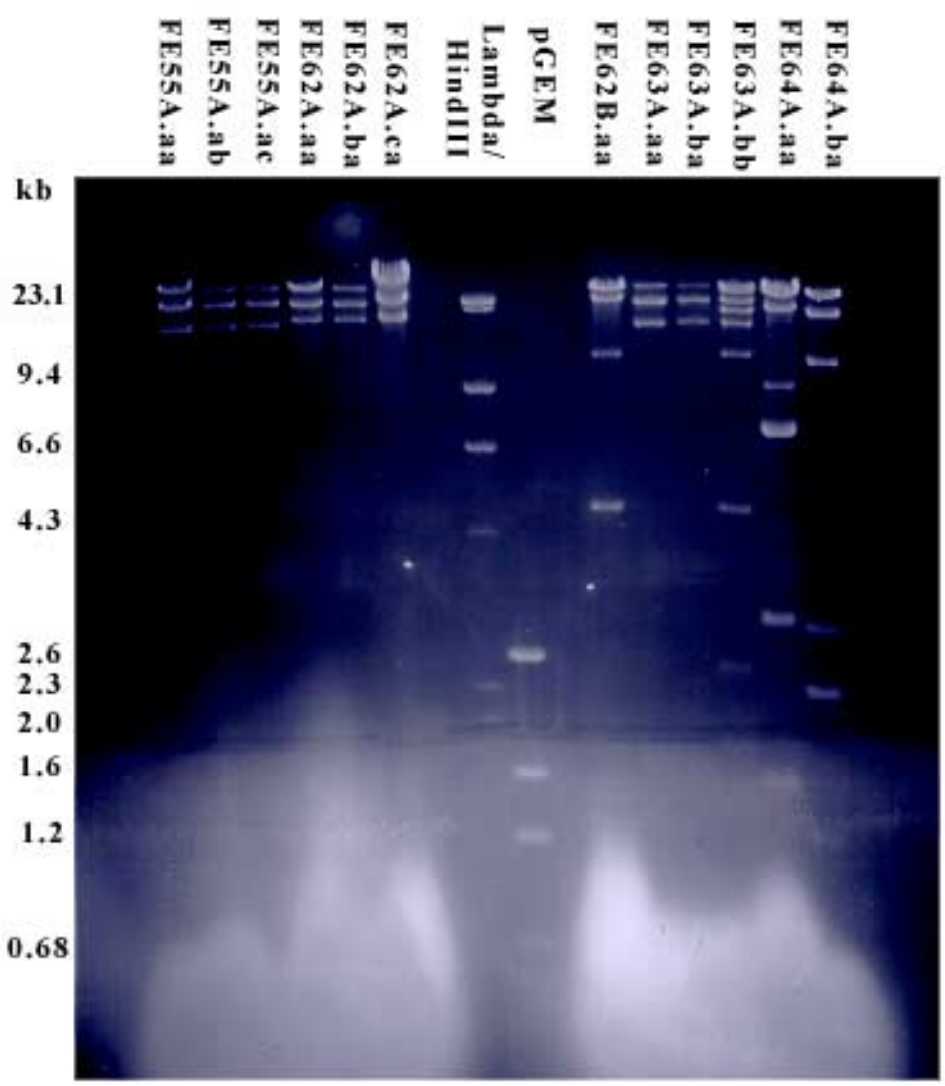
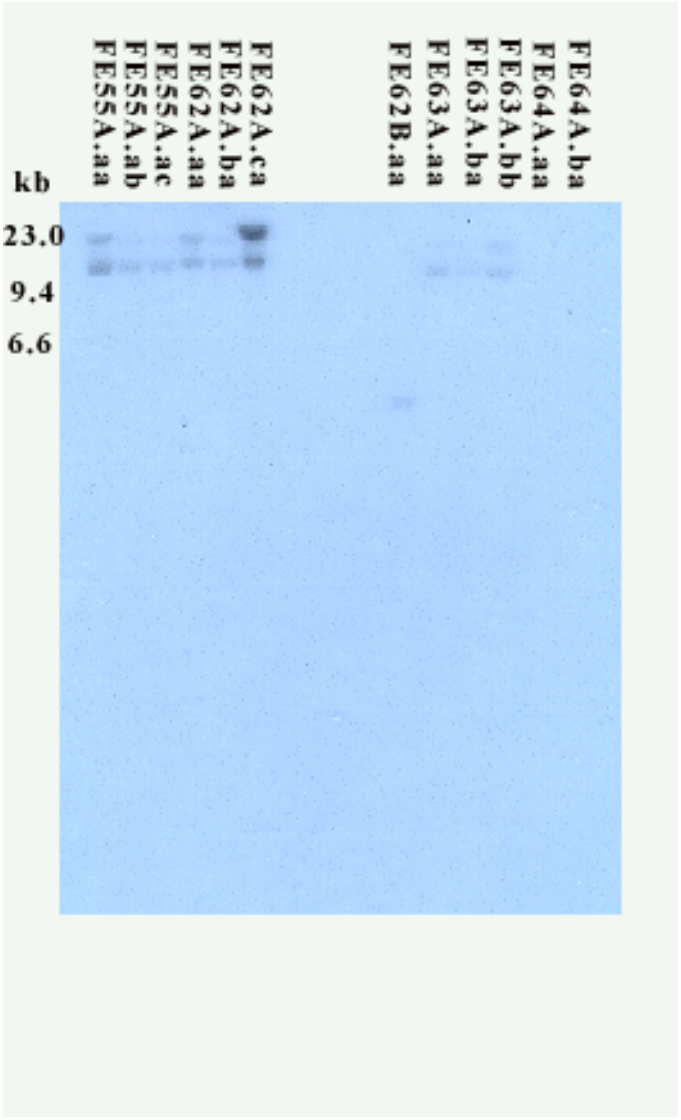

FE55A.aa
FE55A.ab
FE55A.ac
FE62A.aa
FE62A.ba
FE62A.ca
Lambda/
HindIII
pGEM
FE62B.aa
FE63A.aa
FE63A.ba
FE63A.bb
FE64A.aa
FE64A.ba
kb
23.1
9.4
6.6
4.3
2.6
2.3
2.0
1.6
1.2
0.68

Figure 6. Alkaline blot hybridization analysis of *FAD2* genomic clones harbored in the phage EMBL3. The 12 genomic *FAD2* clones (same as in Figure 5) were digested with *Bam*HI, fractionated on an 0.8% agarose gel, and transferred to a positively-charged nylon membrane. The DNA fragments immobilized on the membrane were hybridized with a ^{32}P-labeled probe generated from the 0.8-kb *Xho*I/*Eco*RI fragment of the cotton cDNA clone pSKcF106A at 55˚C. The designation FE shown in the figure for the clones was later changed to LCF.


kb
23.0
9.4
6.6
FE55A.aa
FE55A.ab
FE55A.ac
FE62A.aa
FE62A.ba
FE62A.ca
FE62B.aa
FE63A.aa
FE63A.ba
FE63A.bb
FE64A.aa
FE64A.ba

The genomic clones designated LCFg55A.aa, LCFG62B.aa, and LCFG63A.bb were characterized by agarose gel electrophoresis and alkaline blot hybridization. The sizes determined for the presumptive cotton genomic *FAD2* DNA inserts of LCFg55A.aa, LCFg62B.aa, and LCFg63A.bb were 13.2 kb, 25.0 kb, and 27.0 kb, respectively. The LCFg55A.aa clone was arbitrarily chosen for further physical mapping, and the name was shortened to LCFg55 for convenience. As shown in the physical map of LCFg55 in Figure 7, the 13.2-kb cotton DNA segment containing a putative *FAD2* gene is located between the 20.0-kb lambda left arm and the 9.6-kb lambda right arm. A 10.0-kb *Hin*dIII fragment of LCFg55 DNA was chosen to be subcloned into the phagemid vector pDELTA2 (Gibco BRL Life Technologies) for DNA sequence analysis. The 10.0-kb *Hin*dIII fragment contained a 5.2-kb segment of cotton DNA and a 4.8-kb section of the EMBL3 vector right arm. This *Hin*dIII segment was selected because there were few appropriate hexanucleotide restriction enzyme cleavage sites in the vicinity of the *FAD2* gene compatible with the limited number of restriction sites in the multiple cloning site of the vector pDELTA2. The recombinant plasmid with the 10.0-kb *Hin*dIII fragment of LCFg55 was subcloned into the cosmid vector pDELTA2. The recombinant plasmid designated pCFg55 is shown in the lower panel of Figure 7. The open reading frame of the *FAD2* coding region is located adjacent to the lambda right arm of EMBL3.

Figures 8 and 9 show examples of agarose gel electrophoresis patterns with single and double restriction endonuclease digestions of LCFg55 DNA. Numerous

Figure 7. Physical map of the cotton genomic clone LCFg55 harboring a fatty acid desaturase (*FAD2*) gene. The upper panel shows the 13.2-kb cotton genomic insert (black horizontal line) between the lambda left and lambda right arms of EMBL3 (hatched boxes). The lower panel shows the 10.0-kb *Hin*dIII fragment of LCFg55 subcloned into the cosmid vector pDELTA2 and named pCFg55. The black line represents the 5.2-kb segment of cotton genomic DNA. The black rectangle indicates the 1.2-kb of *FAD2* coding region with the relative polarity of the gene shown from 5′ to 3′. The cross-hatched boxes represent the cosmid vector pDELTA2.

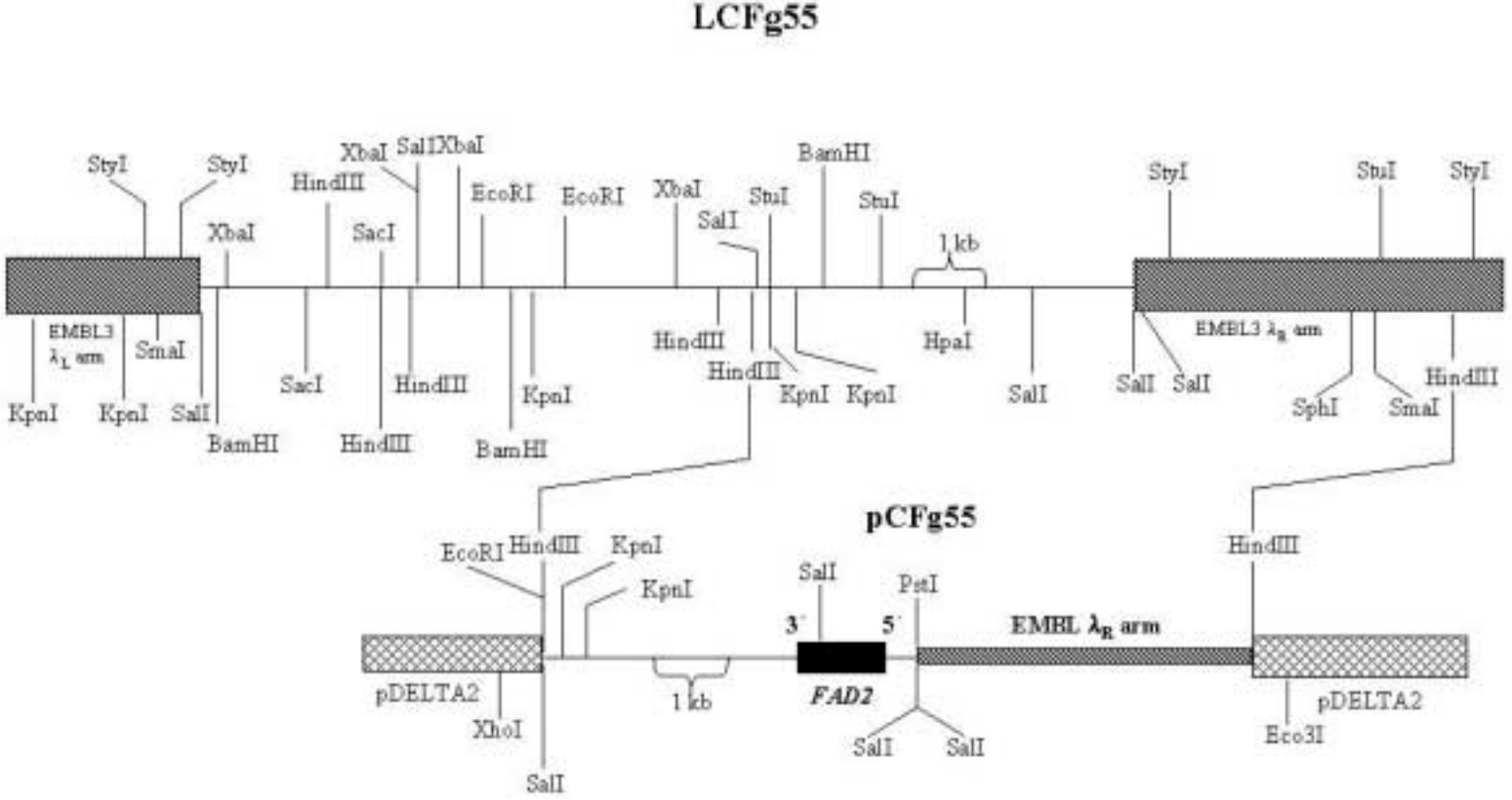

LCFg55
StyI
StyI
XbaI
HindIII
XbaI
SalI
XbaI
EcoRI
EcoRI
XbaI
SalI
StuI
BamHI
StuI
1 kb
StyI
StuI
StyI
EMBL3 λL arm
SmaI
SacI
HindIII
KpnI
BamHI
HindIII
HindIII
KpnI
KpnI
HpaI
SalI
EMBL3 λR arm
SalI
SalI
SphI
SmaI
HindIII
KpnI
KpnI
SalI
BamHI
SacI
HindIII
BamHI
pCFg55
EcoRI
HindIII
KpnI
KpnI
SalI
PstI
HindIII
3'
5'
EMBL λR arm
pDELTA2
1 kb
FAD2
pDELTA2
XhoI
SalI
SalI
SalI
Eco3I
SalI

Figure 8. Electrophoresis of restriction endonuclease fragments generated by digestion of the cotton genomic clone LCFg55 DNA on a 1% agarose gel. The lambda LCFg55 DNA (200 ng per reaction) was digested with restriction endonucleases, fractionated on a 1% agarose gel, and stained with ethidium bromide prior to photography. Standard molecular weight markers (1 μg) are shown in the Lambda/*Hin*dIII, and pGEM lanes. The sizes of these markers are shown on the left.

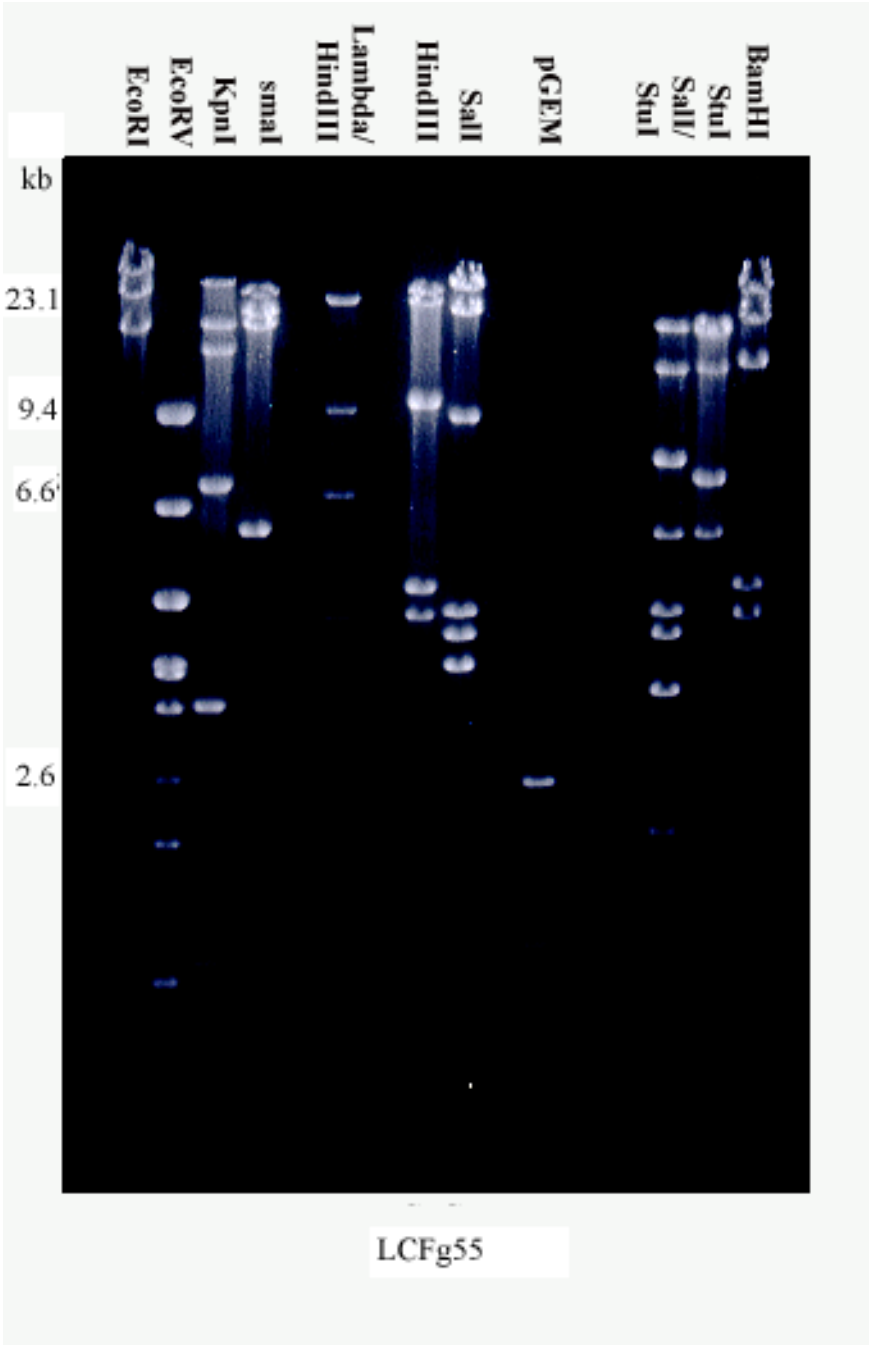
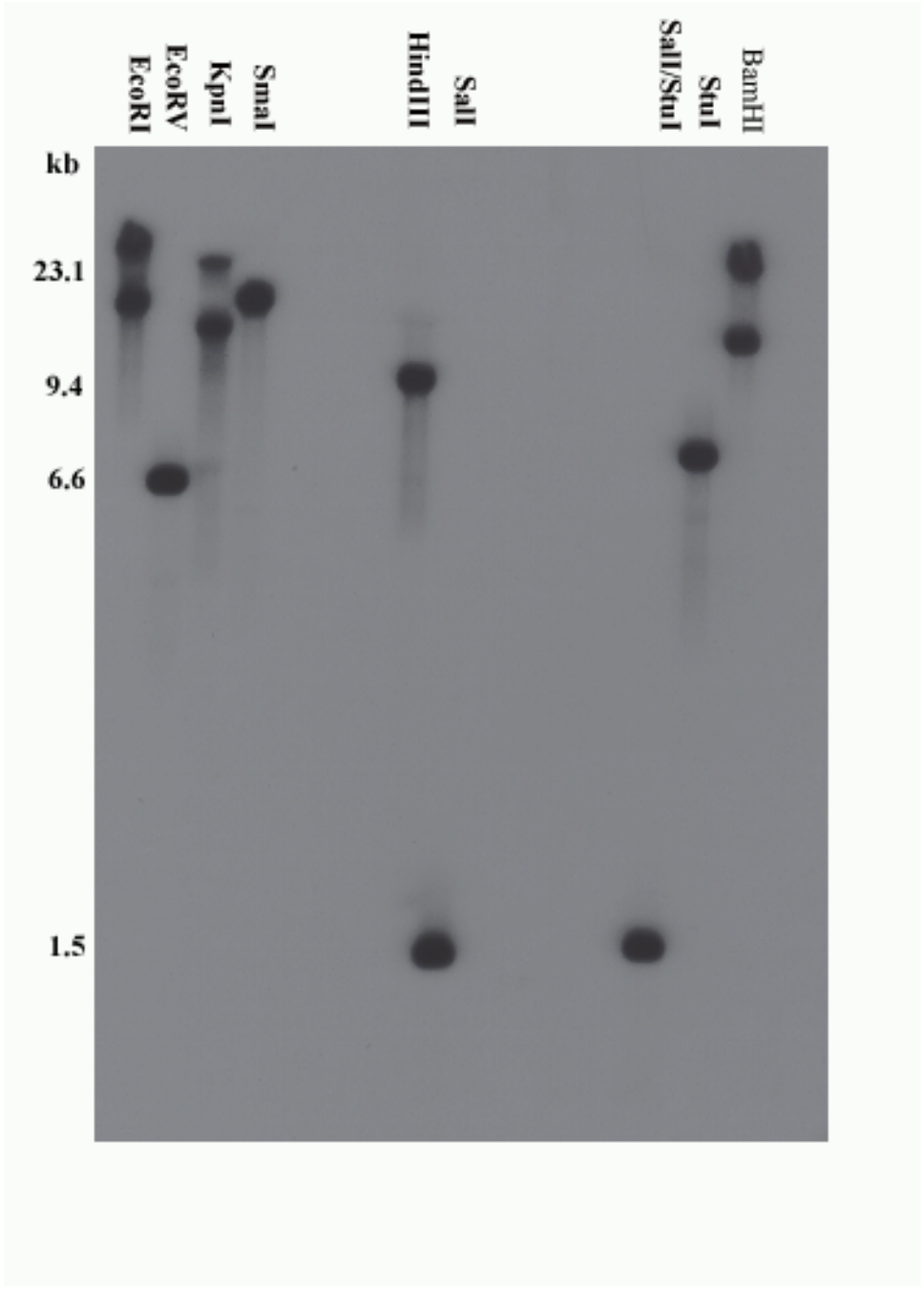

EcoRI
EcoRV
KpnI
smaI
Lambda/
HindIII
SalI
HindIII
pGEM
StuI
SalI/
StuI
BamHI
kb
23.1
9.4
6.6
2.6
LCFg55

Figure 9. Autoradiogram of alkaline blot hybridization of an agarose gel with restriction endonuclease digests of cotton genomic clone LCFg55 DNA. The different restriction endonuclease digestions of LCFg55 DNA were fractionated on a 1% agarose gel, transferred to a positively-charged nylon membrane, and hybridized with a ^{32}P-labeled probe derived from the 0.8-kb *Xho*I/*Eco*RI fragment excised from the cotton *FAD2* cDNA clone pSKcF106A. The two lanes in the middle show the 10.0-kb *Hin*dIII fragment of LCFg55 DNA which was subcloned into the pDELTA2 cosmid vector, and the 1.5-kb *Sal*I fragment of LCFg55, both encompassing the *FAD2* coding region.


EcoRI
EcoRV
KpnI
SmaI
HindIII
SalI
SalI/StuI
StuI
BamHI
kb
23.1
9.4
6.6
1.5

agarose gel patterns were used to determine the sizes of the fragments and subsequently construct the physical map of LCFg55. Upon *Sal*I digestion, it was found that the cotton genomic clone LCFg55 contained a cotton DNA insert of about 13.2 kb. Also, a 1.5-kb *Sal*I DNA fragment that gave an intense signal on the autoradiogram obtained with the ^{32}P-labeled probe derived from the 0.8-kb *Xho*I/*Eco*RI fragment contained the 5′-portion of the *FAD2* coding region.

The structure of the *FAD2* gene in the T-DNA insert of *Arabidopsis* (Okuley et al., 1994) has a 1.6-kb 5′-flanking region with the promoter region and a 1,134-bp intron, only several nucleotides from the ATG start codon. That the 13.2-kb cotton genomic insert in LCFg55 (Figure 7) lacked sufficient 5′-untranslated region to contain both the *FAD2* promoter elements and a large intron became evident upon comparison with the structure of the *Arabidopsis FAD2* gene .

II. Identification of a cotton ω-6 fatty acid desaturase

As previously described, a 10.0-kb *Hin*dIII fragment of the lambda clone LCFg55 was inserted into the cosmid vector pDELTA2 to be used to generate nested deletion subclones by intracellular-based transposon deletion subcloning (GIBCO BRL Life Technologies Deletion Factory 2.0 Manual). The deletion plasmid subclones generated by the transposon (*Tn*1000) are shown in Figure 10. Each deletion subclone was characterized by its relative size on an agarose gel. The deletion map of pCFg55 (Figure 11) revealed that a number of overlapping deletion subclones were generated (e.g. A4, A49, A88, A191, K11). The DNA sequences of the various subclones were

Figure 10. Agarose gel electrophoresis of deletion plasmid subclones generated from the original construct pCFg55. The deletion plasmid DNAs (2 μg per reaction) were digested with *Eco*RI and fractionated on a 0.8% agarose gels. The DNA fragments were stained with ethidium bromide prior to photography. The first lane shows the *Eco*RI-digested original plasmid pCFg55 DNA designated pCFgH3 in this figure (18 kb) which was not subjected to the nested deletion procedure. The subclones designated A2, A47, A49, A52, A96, A127, and A131 are the clockwise deletion subclones of pCFg55. Their sizes were deduced to be 13.0, 11.0, 16.0, 16.0, 10.0, 8.0 and 8.5 kb, respectively, compared with the standard molecular weight marker *Hin*dIII-digested lambda DNA fragments shown on the left (in kb).

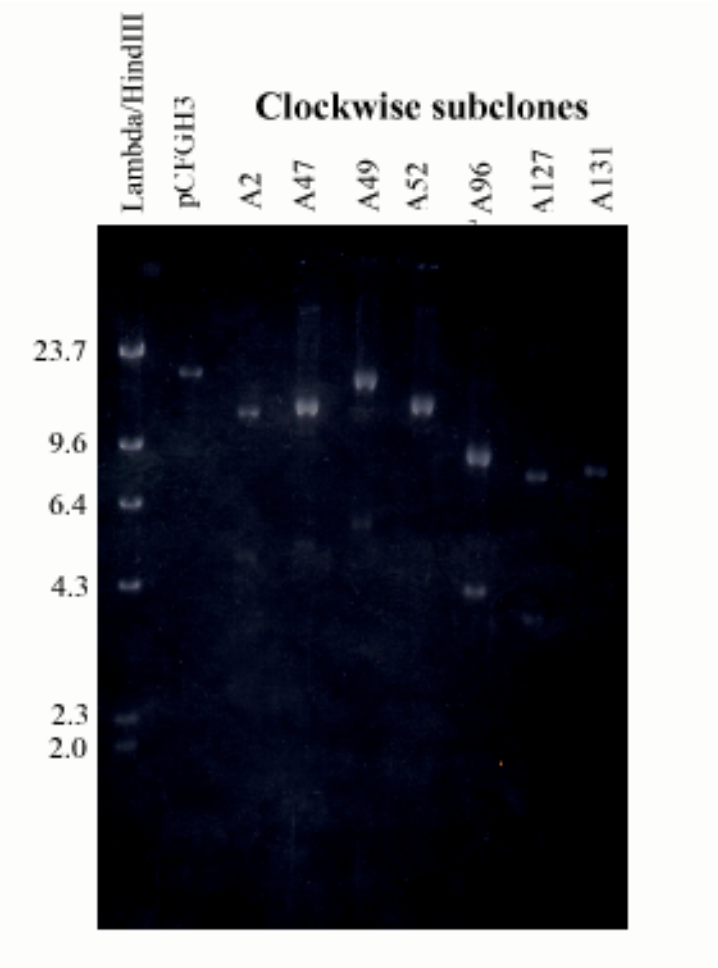
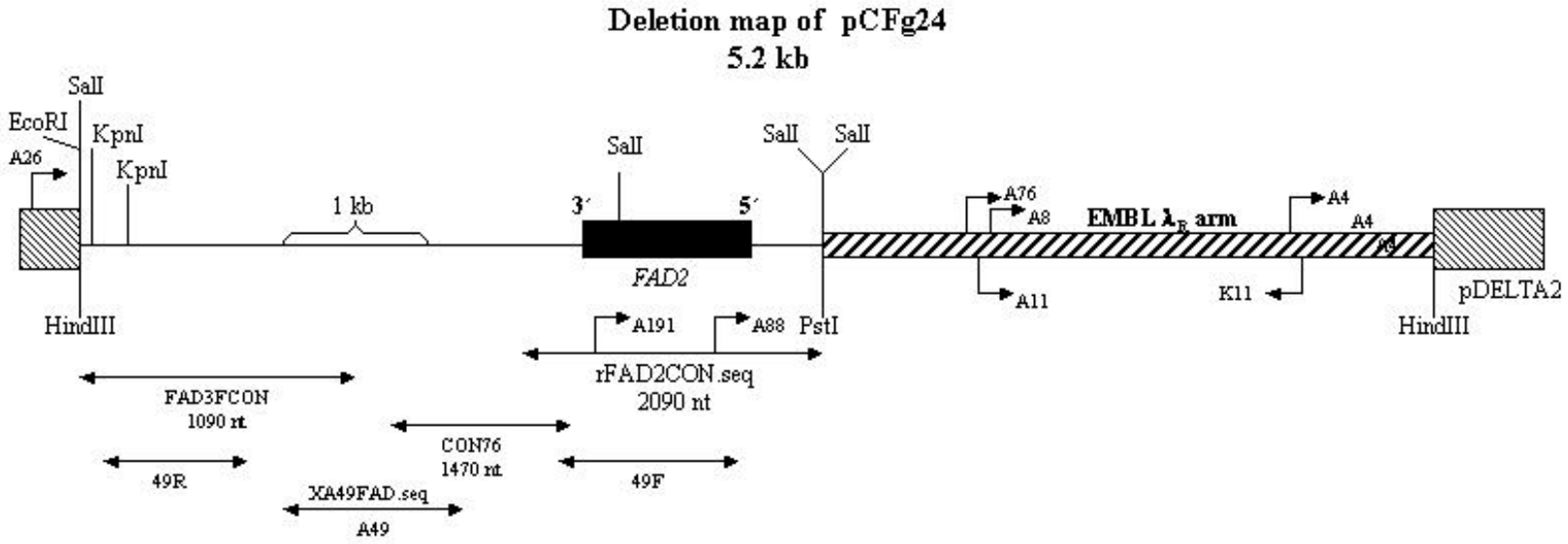

Clockwise subclones
Lambda/HindIII
pCFGH3
A2
A47
A49
A52
A96
A127
A131
23.7
9.6
6.4
4.3
2.3
2.0

Figure 11. Deletion subclone map of the pCFg55 plasmid construct in pDELTA2. A number of clockwise (labeled A) and counterclockwise (labeled K) subclones were generated by transposon-based-nested deletion subcloning (Gibco BRL Life Technologies Deletion Factory 2.0 Manual). A number of overlapping deletion subclones were found (e.g. designated A4, A8, A11, A26, A49, A76, A88, A191, K11).

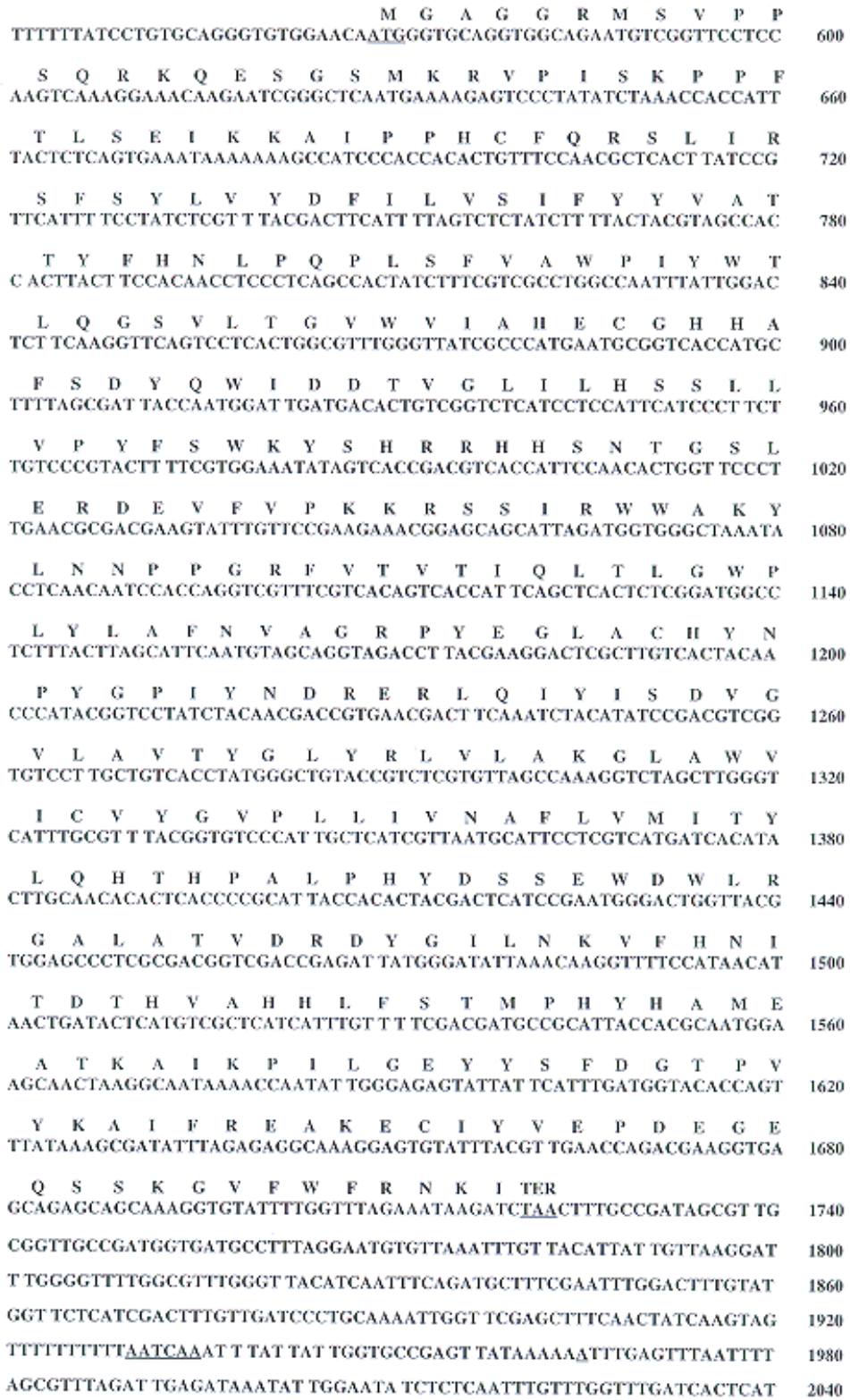

Deletion map of pCFg24
5.2 kb
SalI
EcoRI
KpnI
A26
KpnI
SalI
SalI
SalI
1 kb
3´
5´
A76
A8
EMBL λR arm
A4
A4
FAD2
A11
K11
pDELTA2
HindIII
A191
A88
PstI
HindIII
rFAD2CON.seq
2090 nt
FAD3FCON
1090 nt
CON76
1470 nt
49R
49F
XA49FAD.seq
A49

analyzed by Dr. John Knesek at Texas Woman's University. Some non-overlapping stretches of DNA sequence between the deletion subclones and the sequences of some GC-rich areas with compression effects were manually sequenced by Dr. Irma Pirtle of this laboratory. In addition, the deletion subclones were screened by manual sequencing to avoid sequencing duplicate subclones. As a result, both strands of a 5,216-basepair segment of cotton genomic DNA in the subclone pCFg55 encompassing the *FAD2* gene were sequenced using a combination of a primer-based approach and a nested deletion subclone approach.

The coding region of this cotton *FAD2* gene is 1,155 basepairs, including the start and stop codons (Figure 12). The DNA sequence of the cotton *FAD2* gene has been deposited in GenBank (Accession Number AF331163). The open reading frame encodes a polypeptide of 384 amino acids, and is continuous without introns. This open reading frame has a 5′-flanking region of 568 basepairs (not shown) and a 3′-flanking region of 242 basepairs (Figure 12). A hexanucleotide near-upstream element (AATCAA, underlined at nt 1931 in Figure 12) is thought to be a polyadenylation recognition signal. The putative 3′- polyadenylation site would occur 30 basepairs upstream at nt 1965 (underlined in Figure 12).

The deduced amino acid sequence of the cotton FAD2 polypeptide has significant identity with the amino acid sequence of the *Arabidopsis* FAD2 polypeptide (76% identity on the amino acid sequence level). Therefore, the cotton *FAD2* gene most likely encodes a microsomal ω–6 desaturase. The *Arabidopsis FAD2* gene has a large 1.1-kb intron in its 5′-flanking region, four basepairs upstream from the ATG.

Figure 12. DNA sequence of the coding region of the cotton ω–6 fatty acid desaturase (*FAD2*) gene. The translation initiation codon ATG at nt 572 and the termination codon TAA at nt 1,720 are underlined. The *FAD2* open reading frame has 1,155 basepairs encoding 384 amino acids. The deduced amino acids are represented above the DNA sequence in single-letter code. The putative polyadenylation signal (AATCAA) is located at nt 1,951 (underlined).

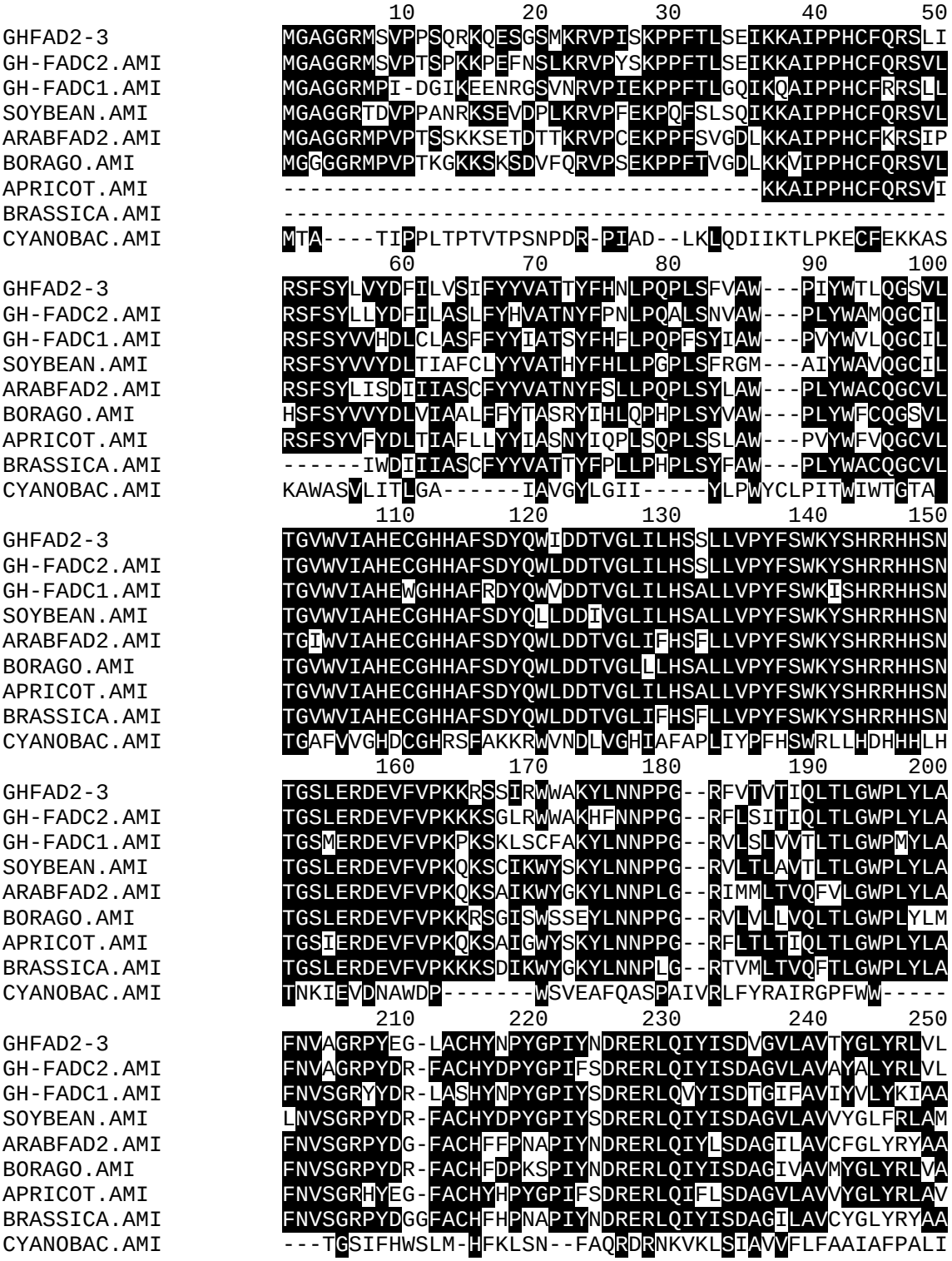

M G A G G R M S V P P
TTTTTTATCCTGTGCAGGGTGTGGAACAATGGGTGCAGGTGGCAGAATGTCGGTTCCTCC 600
12 S Q R K Q E S G S M K R V P I S K P P F
AAGTCAAAGGAAACAAGAATCGGGCTCAATGAAAAGAGTCCCTATATCTAAACCACCATT 660
32 T L S E I K K A I P P H C F Q R S L I R
TACTCTCAGTGAAATAAAAAAAAGCCATCCCACCACACTGTTTCCAACGCTCACT TATCCG 720
52 S F S Y L V Y D F I L V S I F Y Y V A T
TTCATTT TCCTATCTCGT T TACGACTTCATT TTAGTCTCTATCTT TTACTACGTAGCCAC 780
72 T Y F H N L P Q P L S F V A W P I Y W T
C ACTTACT TCCACAACCTCCCTCAGCCACTATCTTTCGTCGCCTGGCCAATTTATTGGAC 840
92 L Q G S V L T G V W V I A H E C G H H A
TCT TCAAGGTTCAGTCCTCACTGGCGTTTGGGTTATCGCCCATGAATGCGGTCACCATGC 900
112 F S D Y Q W I D D T V G L I L H S S L L
TTTTAGCGAT TACCAATGGAT TGATGACACTGTCGGTCTCATCCTCCATTCATCCCT TCT 960
132 V P Y F S W K Y S H R R H H S N T G S L
TGTCCCGTACTT TTCGTGGAAATATAGTCACCGACGTCACCATTCCAACACTGGT TCCCT 1020
152 E R D E V F V P K K R S S I R W W A K Y
TGAACGCGACGAAGTATTTGTTCCGAAGAAACGGAGCAGCATTAGATGGTGGGCTAAATA 1080
172 L N N P P G R F V T V T I Q L T L G W P
CCTCAACAATCCACCAGGTCGTTTCGTCACAGTCACCAT TCAGCTCACTCTCGGATGGCC 1140
192 L Y L A F N V A G R P Y E G L A C H Y N
TCTTTACTTAGCATTCAATGTAGCAGGTAGACCT TACGAAGGACTCGCTTGTCACTACAA 1200
212 P Y G P I Y N D R E R L Q I Y I S D V G
CCCATACGGTCCTATCTACAACGACCGTGAACGACT TCAAATCTACATATCCGACGTCGG 1260
232 V L A V T Y G L Y R L V L A K G L A W V
TGTCCT TGCTGTCACCTATGGGCTGTACCGTCTCGTGTTAGCCAAAGGTCTAGCTTGGGT 1320
252 I C V Y G V P L L I V N A F L V M I T Y
CATTTGCGT T TACGGTGTCCCAT TGCTCATCGTTAATGCATTCCTCGTCATGATCACATA 1380
272 L Q H T H P A L P H Y D S S E W D W L R
CTTGCAACACACTCACCCCGCAT TACCACACTACGACTCATCCGAATGGGACTGGTTACG 1440
292 G A L A T V D R D Y G I L N K V F H N I
TGGAGCCCTCGCGACGGTCGACCGAGAT TATGGGATATTAAACAAGGTTTTCCATAACAT 1500
312 T D T H V A H H I F S T M P H Y H A M E
AACTGATACTCATGTCGCTCATCATTTGT T T TCGACGATGCCGCATTACCACGCAATGGA 1560
332 A T K A I K P I L G E Y Y S F D G T P V
AGCAACTAAGGCAATAAAACCAATAT TGGGAGAGTATTAT TCATTTGATGGTACACCAGT 1620
352 Y K A I F R E A K E C I Y V E P D E G E
TTATAAAGCGATATTTAGAGAGGCAAAGGAGTGTATTTACGT TGAACCAGACGAAGGTGA 1680
372 Q S S K G V F W F R N K I TER
GCAGAGCAGCAAAGGTGTATTTTGGTTTAGAAATAAGATCTAACTTTGCCGATAGCGT TG 1740
CGGTTGCCGATGGTGATGCCTTTAGGAATGTGTTAAATTTGT TACATTAT TGTTAAGGAT 1800
T TGGGGTTTTGGCGTTTGGGT TACATCAATTTCAGATGCTTTCGAATTTGGACTTTGTAT 1860
GGT TCTCATCGACTTTGTTGATCCCTGCAAAATTGGT TCGAGCTTTCAACTATCAAGTAG 1920
TTTTTTTTTTAATCAAAT T TAT TAT TGGTGCCGAGT TATAAAAAATTTGAGTTTAATTTT 1980
AGCGTTTAGAT TGAGATAAATAT TGGAATA TCTCTCAATTTGTTTGGTTTGATCACTCAT 2040

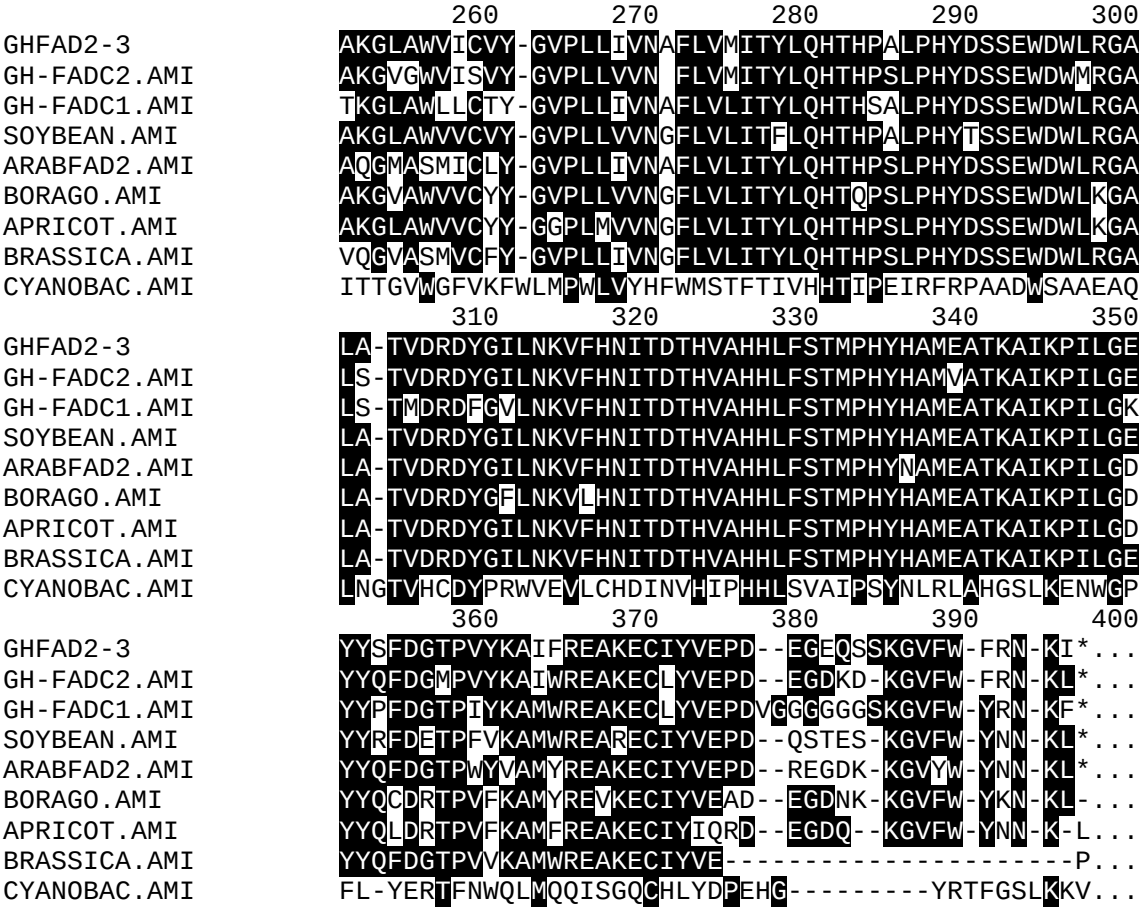
start codon. Thus, it was not surprising that the cotton gene might also have an intron in its 5′-flanking region. However, the coding region of the *FAD2* gene in LCFg55 was near the EMBL3 right vector arm, and the 5′-flanking region in LCFg55 was only 568 basepairs from the ATG start codon (as shown in Figure 7). Therefore, the 5′-untranslated region of the *FAD2* gene was determined to contain only a portion of a potential intron in the 5′-flanking region by comparison with the intron flanking the *Arabidopsis FAD2* gene (Okuley et al., 1994). This cotton genomic clone LCFg55 also lacks the promoter/enhancer elements flanking the *FAD2* gene.

III. Amino acid analysis of the *FAD2* coding region

Comparison of alignments of the deduced amino acid sequences of the cotton FAD2 polypeptide with other plant oleate desaturases was done using DNASIS software (Hitachi) and is shown in Figure 13. The conserved amino acid sequence of the cotton FAD2 shares 85% and 74% identities with the deduced amino acid sequences of the cotton *FAD2-2* and *FAD2-1* cDNA clone, respectively (Liu, et al., 1997; Liu et al., 1999). Thus, this FAD2 sequence was assumed to be a new member of the microsomal ω-6 desaturase gene family in cotton. Also, the sequence similarities between the cotton FAD2 protein and the soybean, *Arabidopsis*, and *Borago* FAD2 proteins is about 75%. Therefore, the FAD2 enzyme analyzed in this dissertation has considerable amino acid sequence similarity with other plant FAD2 enzymes. The cotton FAD2 protein (this dissertation) has a higher identity to the cotton FAD2-2 protein (Liu et al., 1999) than to the cotton FAD2-1 protein (Liu et al., 1997). Compared to this FAD2 polypeptide, lower

Figure 13. Comparison of the deduced amino acid sequences of cotton and other plant FAD2 proteins. These amino acid sequences were aligned using the DNASIS software (Hitachi). The identical amino acid residues are shown by reverse contrast (black and white background) using the standard one-letter amino acid abbreviations. The designations GHFAD2-3.AMI, GH-FADC2.AMI, GH-FADC1.AMI, SOYBEAN.AMI, ARABFAD2.AMI, BORAGO.AMI, APRICOT.AMI, BRASSICA.AMI, and CYANOBAC.AMI refer to the deduced amino acid sequences deduced from the *G. hirsutum FAD2* gene (this dissertation; GenBank AF331163); *G. hirsutum* FAD2-2 cDNA (GenBank Y10112); *G. hirsutum* FAD2-1 cDNA (GenBank X97016); *Glycine max* FAD2-1 cDNA (GenBank L43920); *Arabidopsis thaliana* FAD2 cDNA (GenBank L26296); *Borago officialis* FAD2 cDNA (GenBank AF074324); *Prunus armenica* 06FAD cDNA (GenBank AF071892); *Brassica rapa* FAD2 cDNA (GenBank G4074324); and *Synechocystis* PCC6803 Δ12-desaturase gene (GenBank P20388), respectively.


10 20 30 40 50
GHFAD2-3 MGAGGRMSVPPSQRKQESGSMKRVPISKPPFTLSEIKKAIPPHCFQRSLI
GH-FADC2.AMI MGAGGRMSVPTSPKKPEFNSLKRVPYSKPPFTLSEIKKAIPPHCFQRSVL
GH-FADC1.AMI MGAGGRMPI-DGIKEENRGSVNRVPIEKPPFTLGQIKQAIPPHCFRRSLL
SOYBEAN.AMI MGAGGRTDVPPANRKSEVDPLKRVPFEKPQFSLSQIKKAIPPHCFQRSVL
ARABFAD2.AMI MGAGGRMPVPTSSKKSETDTTKRVPCEKPPFSVGDLKKAIPPHCFKRSIP
BORAGO.AMI MGGGGRMPVPTKGKKSKSDVFQRVPSEKPPFTVGDLKKVIPPHCFQRSVL
APRICOT.AMI ----------KKAIPPHCFQRSVI
BRASSICA.AMI ----------
CYANOBAC.AMI MTA----TIPPLTPTVTPSNPDR-PIAD--LKLQDIIKTLPKECFEKKAS
60 70 80 90 100
GHFAD2-3 RSFSYLVYDFILVSIFYYVATTYFHNLPQPLSFVAW---PIYWTLQGSVL
GH-FADC2.AMI RSFSYLLYDFILASLFYHVATNYFPNLPQALSNVAW---PLYWAMQGCIL
GH-FADC1.AMI RSFSYVVHDLCLASFFYYIATSYFHFLPQPFSYIAW---PVYWVLQGCIL
SOYBEAN.AMI RSFSYVVYDLTIAFCLYYVATHYFHLLPGPLSFRGM---AIYWAVQGCIL
ARABFAD2.AMI RSFSYLISDIIIASCFYYVATNYFSLLPQPLSYLAW---PLYWACQGCVL
BORAGO.AMI HSFSYVVYDLVIAALFFYTASRYIHLQPHPLSYVAW---PLYWFCQGSVL
APRICOT.AMI RSFSYVFYDLTIAFLLYYIASNYIQPLSQPLSSLAW---PVYWFVQGCVL
BRASSICA.AMI ------IWDIIIASCFYYVATTYFPLLPHPLSYFAW---PLYWACQGCVL
CYANOBAC.AMI KAWASVLITLGA------IAVGYLGII-----YLPWYCLPITWIWTGTA
110 120 130 140 150
GHFAD2-3 TGVWVIAHECGHHAFSDYQWIDDTVGLILHSSLLVPYFSWKYSHRRHHSN
GH-FADC2.AMI TGVWVIAHECGHHAFSDYQWLDDTVGLILHSSLLVPYFSWKYSHRRHHSN
GH-FADC1.AMI TGVWVIAHEWGHHAFRDYQWVDDTVGLILHSALLVPYFSWKISHRRHHSN
SOYBEAN.AMI TGVWVIAHECGHHAFSDYQLLDDIVGLILHSALLVPYFSWKYSHRRHHSN
ARABFAD2.AMI TGIWVIAHECGHHAFSDYQWLDDTVGLIFHSFLLVPYFSWKYSHRRHHSN
BORAGO.AMI TGVWVIAHECGHHAFSDYQWLDDTVGLLLHSALLVPYFSWKYSHRRHHSN
APRICOT.AMI TGVWVIAHECGHHAFSDYQWLDDTVGLILHSALLVPYFSWKYSHRRHHSN
BRASSICA.AMI TGVWVIAHECGHHAFSDYQWLDDTVGLIFHSFLLVPYFSWKYSHRRHHSN
CYANOBAC.AMI TGAFVVGHDCGHRSFAKKRWVNDLVGHIAFAPLIYPFHSWRLLHDHHHLH
160 170 180 190 200
GHFAD2-3 TGSLERDEVFVPKKRSSIRWWAKYLNNPPG--RFVTVTIQLTLGWPLYLA
GH-FADC2.AMI TGSLERDEVFVPKKKSGLRWWAKHFNNPPG--RFLSITIQLTLGWPLYLA
GH-FADC1.AMI TGSMERDEVFVPKPKSKLSCFAKYLNNPPG--RVLSLVVTLTLGWPMYLA
SOYBEAN.AMI TGSLERDEVFVPKQKSCIKWYSKYLNNPPG--RVLTLAVTLTLGWPLYLA
ARABFAD2.AMI TGSLERDEVFVPKQKSAIKWYGKYLNNPLG--RIMMLTVQFVLGWPLYLA
BORAGO.AMI TGSLERDEVFVPKKRSGISWSSEYLNNPPG--RVLVLLVQLTLGWPLYLM
APRICOT.AMI TGSIERDEVFVPKQKSAIGWYSKYLNNPPG--RFLTLTIQLTLGWPLYLA
BRASSICA.AMI TGSLERDEVFVPKKKSDIKWYGKYLNNPLG--RTVMLTVQFTLGWPLYLA
CYANOBAC.AMI TNKIEVDNAWDP-------WSVEAFQASPAIVRLFYRAIRGPFWW-----
210 220 230 240 250
GHFAD2-3 FNVAGRPYEG-LACHYNPYGPIYNDRERLQIYISDVGVLAVTYGLYRLVL
GH-FADC2.AMI FNVAGRPYDR-FACHYDPYGPIFSDRERLQIYISDAGVLAVAYALYRLVL
GH-FADC1.AMI FNVSGRYYDR-LASHYNPYGPIYSDRERLQVYISDTGIFAVIYVLYKIAA
SOYBEAN.AMI LNVSGRPYDR-FACHYDPYGPIYSDRERLQIYISDAGVLAVVYGLFRLAM
ARABFAD2.AMI FNVSGRPYDG-FACHFFPNAPIYNDRERLQIYLSDAGILAVCFGLYRYAA
BORAGO.AMI FNVSGRPYDR-FACHFDPKSPIYNDRERLQIYISDAGIVAVMYGLYRLVA
APRICOT.AMI FNVSGRHYEG-FACHYHPYGPIFSDRERLQIFLSDAGVLAVVYGLYRLAV
BRASSICA.AMI FNVSGRPYDGGFACHFHPNAPIYNDRERLQIYISDAGILAVCYGLYRYAA
CYANOBAC.AMI ---TGSIFHWSLM-HFKLSN--FAQRDRNKVKLSIAVVFLFAAIAFPALI

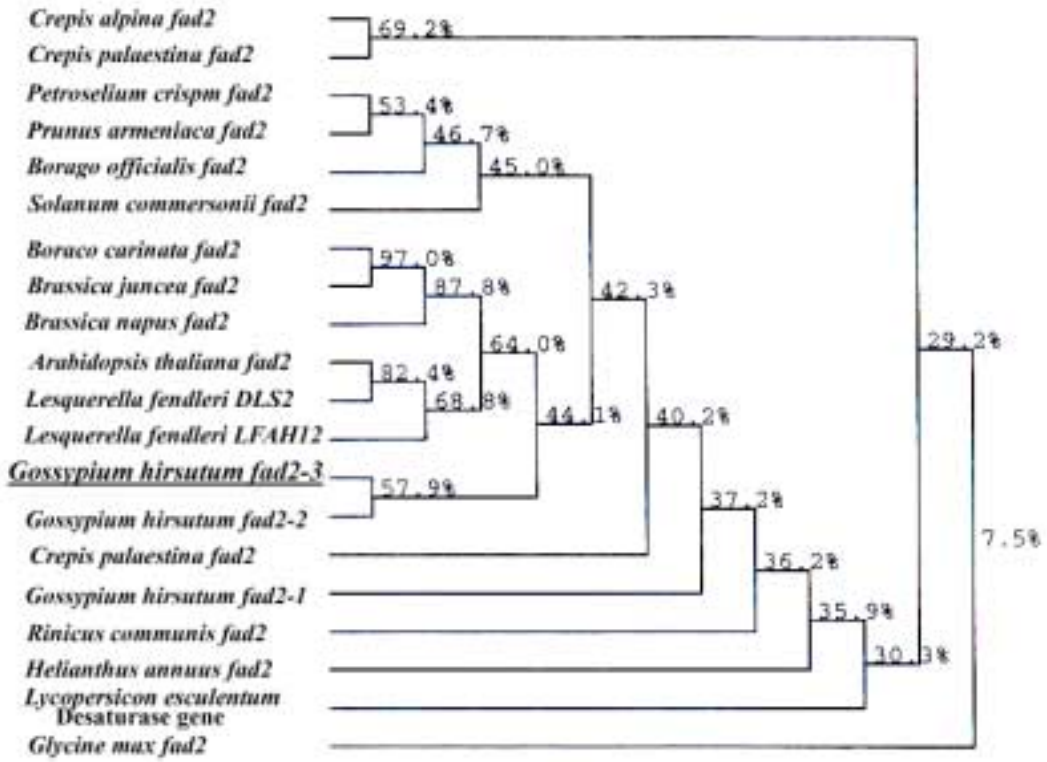

260 270 280 290 300
GHFAD2-3 AKGLAWVICVY-GVPLLIVNAFLVMITYLQHTHPALPHYDSSEWDWLRGA
GH-FADC2.AMI AKGVGWVISVY-GVPLLVVN FLVMITYLQHTHPSLPHYDSSEWDWMRGA
GH-FADC1.AMI TKGLAWLLCTY-GVPLLIVNAFLVLITYLQHTHSALPHYDSSEWDWLRGA
SOYBEAN.AMI AKGLAWVVCVY-GVPLLVVNGFLVLITFLQHTHPALPHYTSSEWDWLRGA
ARABFAD2.AMI AQGMASMICLY-GVPLLIVNAFLVLITYLQHTHPSLPHYDSSEWDWLRGA
BORAGO.AMI AKGVAWVVCYY-GVPLLVVNGFLVLITYLQHTQPSLPHYDSSEWDWLKGA
APRICOT.AMI AKGLAWVVCYY-GGPLMVVNGFLVLITYLQHTHPSLPHYDSSEWDWLKGA
BRASSICA.AMI VQGVASMVCFY-GVPLLIVNGFLVLITYLQHTHPSLPHYDSSEWDWLRGA
CYANOBAC.AMI ITTGVWGFVKFWLMPWLVYHFWMSTFTIVHHTIPEIRFRPAADWSAAEAQ
310 320 330 340 350
GHFAD2-3 LA-TVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGE
GH-FADC2.AMI LS-TVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMVATKAIKPILGE
GH-FADC1.AMI LS-TMDRDFGVLNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGK
SOYBEAN.AMI LA-TVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGE
ARABFAD2.AMI LA-TVDRDYGILNKVFHNITDTHVAHHLFSTMPHYNAMEATKAIKPILGD
BORAGO.AMI LA-TVDRDYGFLNKVLHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGD
APRICOT.AMI LA-TVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGD
BRASSICA.AMI LA-TVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGE
CYANOBAC.AMI LNGTVHCDYPRWVEVLCHDINVHIPHHLSVAIPSYNLRLAHGSLKENWGP
360 370 380 390 400
GHFAD2-3 YYSFDGTPVYKAIFREAKECIYVEPD--EGEQSSKGVFW-FRN-KI*...
GH-FADC2.AMI YYQFDGMPVYKAIWREAKECLYVEPD--EGDKD-KGVFW-FRN-KL*...
GH-FADC1.AMI YYPFDGTPIYKAMWREAKECLYVEPDVGGGGGGSKGVFW-YRN-KF*...
SOYBEAN.AMI YYRFDETPFVKAMWREARECIYVEPD--QSTES-KGVFW-YNN-KL*...
ARABFAD2.AMI YYQFDGTPWYVAMYREAKECIYVEPD--REGDK-KGVYW-YNN-KL*...
BORAGO.AMI YYQCDRTPVFKAMYREVKECIYVEAD--EGDNK-KGVFW-YKN-KL-...
APRICOT.AMI YYQLDRTPVFKAMFREAKECIYIQRD--EGDQ--KGVFW-YNN-K-L...
BRASSICA.AMI YYQFDGTPVVKAMWREAKECIYVE----------P...
CYANOBAC.AMI FL-YERTFNWQLMQQISGQCHLYDPEHG---------YRTFGSLKKV...

amino acid sequence identities were found with the *Prunus armenica* (apricot) FAD protein (72%) and the *Brassica rapa* FAD2 protein (66%). The amino acid sequences of both *P. armenica* FAD and *B. rapa* FAD2 proteins were deduced from their partial cDNA sequences. Thus, they lack about 30-35 N-terminal amino acids when the plant *FAD2* cDNA sequences were compared by dendrogram analysis. The cyanobacterium *Synechocystis* PCC6803 had the lowest identity (28%) among the deduced FAD2 amino acid sequences. The amino acid sequences were deduced from aligned cDNA sequences obtained from the GenBank database and analyzed by DNASIS software.

IV. Phylogenetic relationships between the cotton FAD2 open reading frame with the open reading frames of other membrane-bound fatty-acyl desaturases

The phylogenetic relationships among the DNA sequences of the open reading frames of the plant fatty acyl desaturases are shown by construction of dendrogram plots based on their DNA sequence similarities (Figure 14). The soluble acyl-ACP desaturases are apparently not related to other membrane-bound desaturases (Tocher et al., 1998). This close relationship is restricted to the plastid fatty acyl desaturases in higher plants, and is not found in cyanobacteria (Tocher et al., 1998). The soluble acyl-ACP and membrane-bound desaturases in higher plants and the cyanobacterial Δ12-desaturase apparently use different mechanisms to determine the specific positions of double bond insertion into the acyl group backbone of fatty acids, based on the poor structural relationships in the phylogenetic tree and their predicted primary amino acid structures. The relationship between the cotton *FAD2* gene sequence and microsomal ω-6 desaturase

Figure 14. Dendrogram analysis of *FAD*-like DNA sequences in the plant desaturases. This analysis was done by DNASIS software (Hitachi). The DNA sequences were obtained from the GenBank Database. The length of the branches is proportional to the sequence differences among plant cDNA/gene sequences. The sequence identities are shown as percentages above the horizontal lines. The *Gossypium hirsutum fad2-3* is the *FAD2* sequence described in this dissertation.


Crepis alpina fad2
Crepis palaestina fad2
Petroselium crispm fad2
Prunus armeniaca fad2
Borago officialis fad2
Solanum commersonii fad2
Boraco carinata fad2
Brassica juncea fad2
Brassica napus fad2
Arabidopsis thaliana fad2
Lesquerella fendleri DLS2
Lesquerella fendleri LFAH12
Gossypium hirsutum fad2-3
Gossypium hirsutum fad2-2
Crepis palaestina fad2
Gossypium hirsutum fad2-1
Rinicus communis fad2
Helianthus annuus fad2
Lycopersicon esculentum
Desaturase gene
Glycine max fad2
69.2%
53.4%
46.7%
45.0%
97.0%
87.8%
64.0%
82.4%
68.8%
44.1%
57.9%
42.3%
40.2%
37.2%
36.2%
35.9%
30.3%
29.2%
7.5%

cDNA sequences is shown in Figure 14. The DNA sequences of several plant *FAD2* cDNAs were compared in the dendrogram analysis. The aligned cDNA sequences from the GenBank database were analyzed using DNASIS software (Hitachi), and the degree of sequence identities is shown as percentages. The cotton *FAD2* gene sequence (this dissertation) showed significant similarity with numerous plant microsomal ω-6-desaturase cDNA sequences, including *Arabidopsis FAD2* cDNA (68%) and *Brassica FAD2* cDNA (64%). There is a lower sequence similarity between the cotton *FAD2* gene and the *Gossypium FAD2-1* cDNA (37%), *Rinicus FAD2* cDNA (36.2%), and *Helianthus FAD2* cDNA (35.9%) sequences. The lowest cDNA sequence identity is with that of *Glycine max FAD2* (7.5%). Thus, it is possible that all plant ω-3 desaturases and membrane-bound desaturases may be derived from a common ancestral parent.

V. Hydropathy plot of the cotton FAD2 amino acid sequence

The structures of ω-6 fatty acid desaturases can be divided into two types, either the plastidal type or the microsomal type, inferred from the presence of a plastidial import signal sequence or an endoplasmic reticulum targeting signal (Jackson et al., 1990). Another criterion for differentiating the types of desaturases is the presence of the three conserved histidine motifs (Ohlrogge and Browse, 1995). These clusters might be involved in the formation of the active sites in the structures of desaturases (Los and Murata, 1998; Tocher et al., 1998). The histidine-rich motifs are observed in the sequences of membrane-bound Δ6, Δ12 and ω-3 desaturases in higher plants (Tocher et al., 1998). In addition, bacterial membrane-bound alkane hydroxylase and xylene monooxygenase also contain the catalytic diiron-atom binding sites. These proteins are

also membrane-bound desaturases, since they are probably composed of three conserved histidine domains located on the cytoplasmic face of the membrane and have two long hydrophobic domains spanning the membrane twice (Shankin et al., 1994; Stukey et al, 1990).

A hydropathy plot generated by the method of Kyte and Doolittle (1982) in DNASIS (Hitachi) software was done on the predicted amino acid sequence of the FAD2 polypeptide. As shown in Figure 15, the amino acid sequence of FAD2 is very hydrophobic and hence the cotton FAD2 probably would not be a soluble protein. This is because six potential membrane-spanning domains (numbered) occur in the protein. These domains may represent regions that anchor the polypeptide in the endoplasmic reticulum membrane. Thus, the cotton FAD2 would be an integral trans-membrane protein. In addition, the three histidine-rich motifs (indicated by H and underlined) are found in the amino acid sequence of the cotton FAD2 (Figure 15 and Figure 16). These three histidine-rich motifs are homologous with the other acyl-lipid and acyl-CoA desaturases (Los and Murata, 1998; Shanklin and Cahoon, 1998).

Figure 16 shows the locations of the three conserved histidine-rich motifs (underlined) in the cotton FAD2 amino acid sequence. These predicted histidine motifs are located at amino acids 105 to 112 (HECGHHAF), amino acids 137 to 145 (EXXXXXHXXHH), and amino acids 315 to 322 in the carboxy terminus (HXXHH motif HVAHHLFS). From the hydropathy plot, these three histidine-rich motifs are

Figure 15. Hydropathy plot of the cotton FAD2 sequence. This plot was made by the method of Kyte and Doolittle (Kyte and Doolittle, 1982; Devereux et al., 1984) using DNASIS (Hitashi) software. The six putative membrane-spanning domains are indicated by the numbers 1-6. The three H's represent the three histidine-rich motifs found in the acyl-lipid and acyl-CoA desaturases (Los and Murata, 1998; Shanklin and Cahoon, 1998).

File: GHFAD2-3 1 - 385
Table: Kyte & Doolittle
Window: 10 Average: -0.06 Threshold Line: 0.00

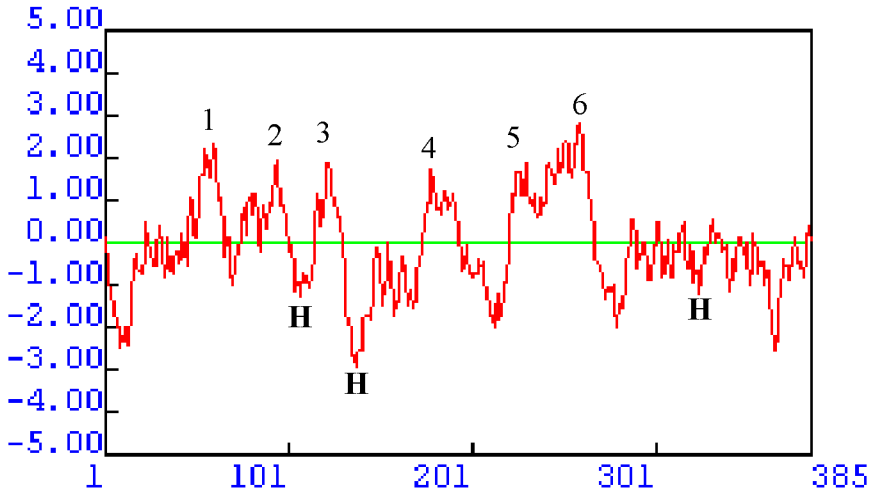


Figure 16. The amino acid sequence showing the conserved histidine motifs of the cotton FAD2 polypeptide in this dissertation. The locations of the three conserved histidine-rich motifs are represented by bold and underlined letters. There are three consensus histidine-rich motifs, HXCGHH (amino acids 105 to 112), EXXXXXHXXHH (amino acids 137 to 145), and HXXHH motifs (amino acids 315 to 322).

10 20 30 40

MGAGGRMSVP PSQRKQESGS MKRVPISKPP FTLSEIKKAI

50 60 70 80

PPHCFQRSLI RSFSKLVYDF ILVSIFYYVA TTYFHNLPQP

90 100 110 120

LSFVAWPIYW TLQGSVLTGV WVIA**HECGHH** AFSDYQWIDD

140 150 160 170

LVPYFSWKYS **HRRHH**NTGS LERDEVFVPK KRSSIRWWAK

180 190 200 210

YLNNPPGRFV TVTIQLTLGW LVPYFSWKYS RPYEGLACHY

220 230 240 250

NYPGPIYNGR ERLQIYISDV GVLAVTYGLY RLVLAKGLAW

260 270 280 290

VICVYGVPLL IVNAFLVMIT YLQHTHPALP HYDSSEWDWL

300 310 320 330

RGALATVDRD YGILNKVFHN ITDT**HVAHH**L FSTMPHYHAM

340 350 360 370

EATKAIKPIL GEYYSFDGTG VYKAIFREAK ECIYVEPDEG

380

EQSSKGVFWF RNKI

present in a similar locations in the eight FAD2 protein sequences listed in the amino acid sequence alignment in Figure 13.

VI. Isolation of the cotton *FAD2* gene in the overlapping genomic clone LCFg24

How the *FAD2* genes of plants are regulated and involved in fatty acid biosynthesis is still unknown. The expression of the cotton *FAD2* gene can be controlled at many levels, such as the initiation of transcription, RNA processing, translation, protein processing, and post-translational modification (Heppard et al., 1996). The promoter, enhancer elements, and transcription factors, as well as the accessory factors, all participate in the initiation of transcription. For example, the GT-1 binding site (Box II and III) and its nuclear factor are important in the light-responsive pea rbc S-3A gene (Green et al., 1988). In addition, RNA processing and splicing of the *FAD2* heterogeneous nuclear RNA (HnRNA) play a role in the synthesis of the mature mRNA. The targeting signals directing cell sorting of membrane proteins to the ER may be important for targeting the FAD2 protein to the ER. The KDEL sequence and the carboxy-terminal tetrapeptide and di-lysine or di-arginine targeting signal may both be necessary and sufficient for the retention of soluble resident proteins in the ER (Teasdale and Jackson, 1996). Furthermore, the mechanisms involved in the production and export of fatty acids from plastids to the rest of the cell may have an influence on the synthesis of the FAD2 polypeptide and its expression (Ohlrogge and Browse, 1995). Another key question in the study of gene regulation in plants is how the *FAD2* gene is expressed in specific cell types and tissues. An analysis of the promoter elements may address this question. Thus, screening for the overlapping *FAD2* clone LCFg24 was

necessary so that the presumptive upstream promoter control elements could be identified for analysis by electrophoresis mobility shift assays and DNA footprinting studies in the near future.

The cotton *FAD2* genomic clone LCFg55 isolated from the library harbored in EMBL3 unfortunately contained too small a genomic cotton DNA insert because it lacks the promoter/enhancer elements when compared to the structure of the *Arabidopsis FAD2* gene. The *FAD2* gene in LCFg55 has a short 5´-untranslated region of about 568 basepairs before the ATG start codon, close to the EMBL3 right arm (Figure 7). The *Arabidopsis FAD2* gene (Okuley et al., 1994) has a large 1.1-kb intron in its 5'-flanking region, an uninterrupted coding region, and about 100 bp of 3´-untranslated region (GenBank L26296). By analogy with the location and size of the intron relative to the coding region of the *Arabidopsis FAD2* gene, the cotton *FAD2* gene in LCFg55 would have only about half of a potential intron in the 5´-flanking region of the LCFg55 DNA. Thus, an overlapping genomic clone of the *FAD2-3* gene had to be isolated by chromosomal walking. The 5´-untranslated region of the *FAD2* gene in LCFg55 was used as a target region to design a unique 5´-flanking region probe for use with PCR to isolate another genomic clone overlapping the *FAD2* gene in LCFg55. As described in Methods, eight genomic clones from the second genomic library were identified with the probe derived from a gene-specific 5′-flanking region template. An alkaline blot hybridization with the 5′-flanking region probe is shown in Figure 18. Five of the eight clones gave different signal intensities with the ^{32}P-labeled probe. As shown in Figures

Figure 17. Agarose gel electrophoresis of restriction enzyme fragments from the eight genomic clones encompassing the putative *FAD2* gene in the lambda FIXII vector. The eight cloned DNAs (1.5 μg per reaction) were digested with *Bam*HI and *Sst*I and fractionated on an 0.8% agarose gel. The DNA fragments were stained with ethidium bromide prior to photography. The genomic clone 13ca1a was not amplified by the minilysate phage preparation. The clones L23ac8a, L23bc1a, L23cd1a, and L24ba1a were judged to have similar cotton DNA inserts. The clones L13ba8a and L24ba2b were also found to have similar cotton DNA inserts. The 6.8-kb *Bam*HI/*Sst*I genomic fragment in L24ba2b was subcloned into pUC19 for further analysis and designated pCFg24H. The size markers of the standard fragments from a *Hin*dIII digest of lambda DNA and from pGEM markers (Promega) are shown at the left in kb.

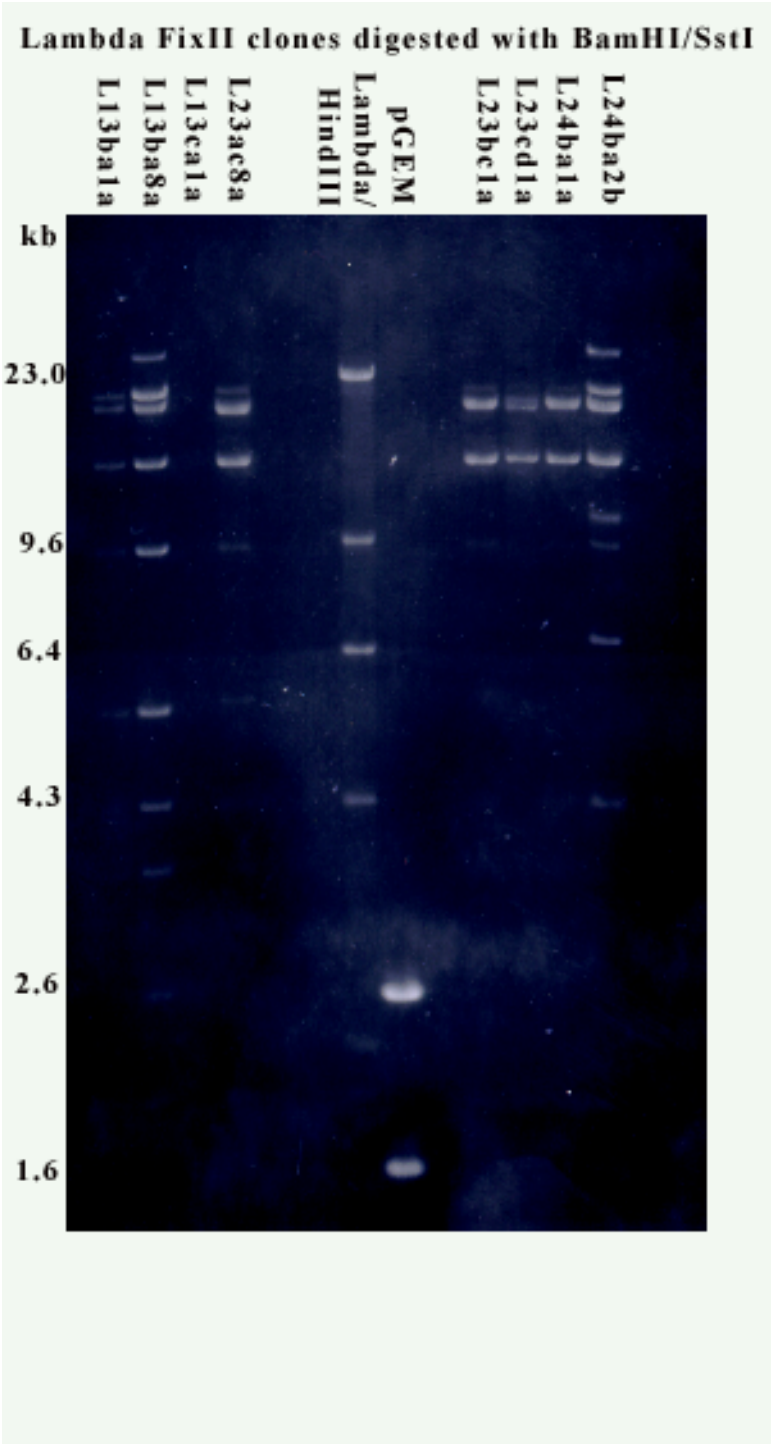
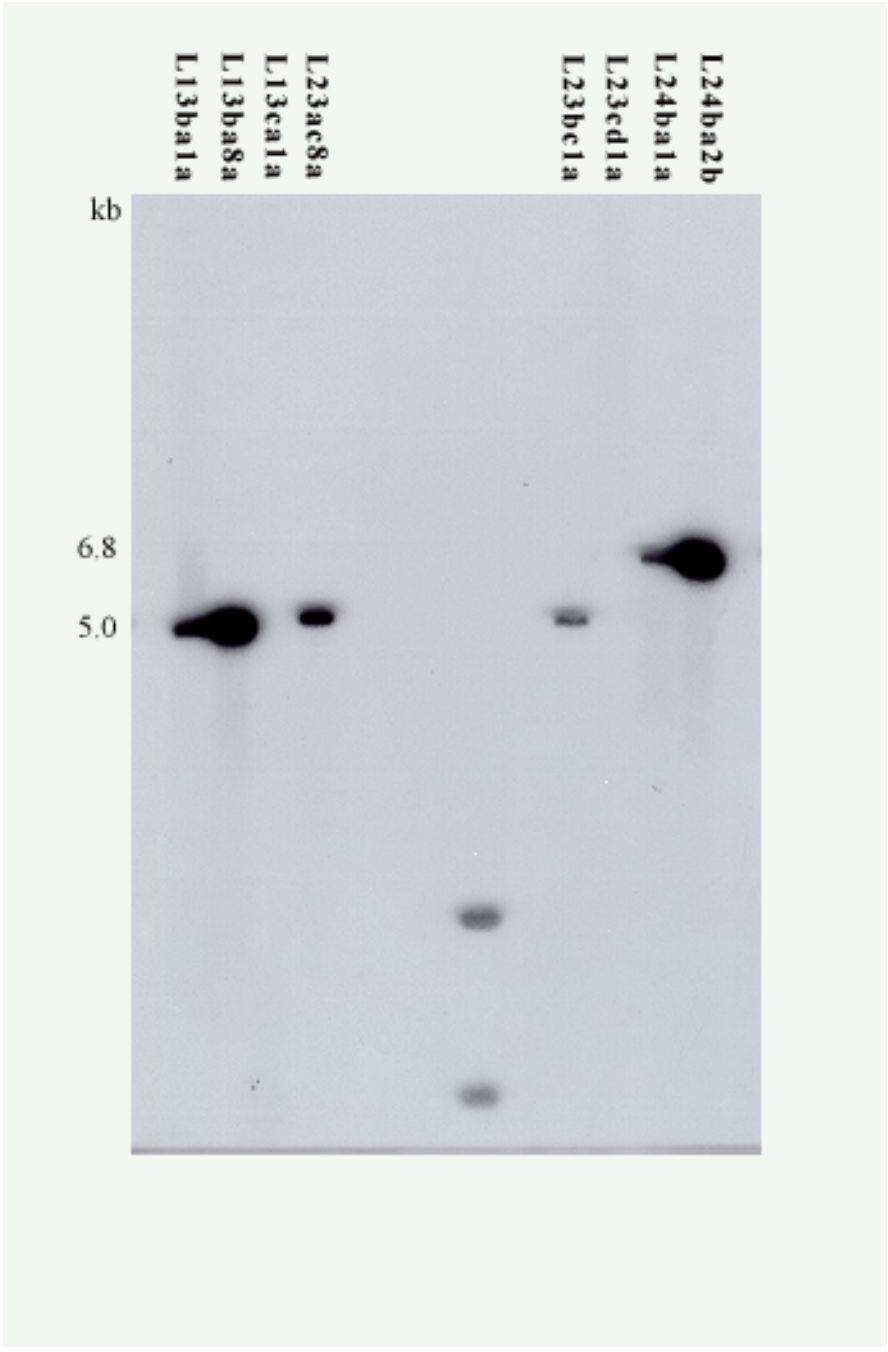

Lambda FixII clones digested with BamHI/SstI
L13ba1a
L13ba8a
L13ca1a
L23ac8a
Lambda/ HindIII
pGEM
L23bc1a
L23cd1a
L24ba1a
L24ba2b
kb
23.0
9.6
6.4
4.3
2.6
1.6

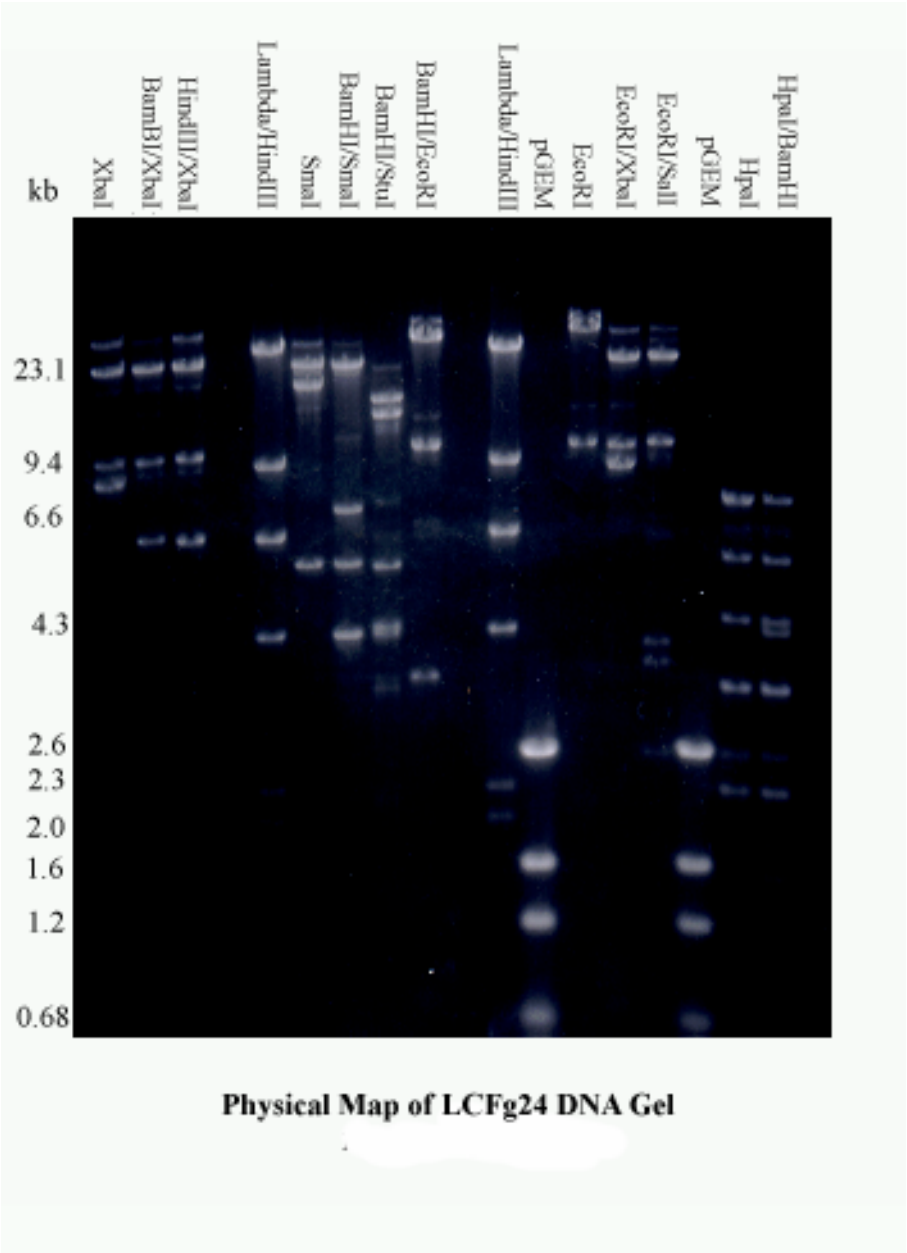
Figure 18. Autoradiogram of an alkaline blot of eight genomic clones encompassing possible *FAD2* genes in the lambda FIXII vector. *Bam*HI/*Sst*I digests of the eight genomic clones were fractionated on a 0.8% agarose gel (Figure 17), transferred to a positively-charged nylon membrane, and hybridized with a ^{32}P-labeled probe derived from the 5′-flanking fragment of the *FAD2* gene. Six of eight of the genomic clones showed an intense signal with the ^{32}P-labeled 5′-flanking region probe. The clones L13ba1a and L24ba2b gave the most intense signal of the 5.0-kb and 6.8-kb *Bam*HI/*Sst*I DNA fragments, respectively. The sizes (in kb) of the 5.0 and 6.8 *Bam*HI/*Sst*I fragments are shown at the left. The two lower hybridizing bands in the middle are from nonspecific hybridization to the pGEM standard markers.


L13ba1a
L13ba8a
L13ca1a
L23ac8a
L23bc1a
L23cd1a
L24ba1a
L24ba2b
kb
6.8
5.0

17 and 18, a 5.0-kb *Bam*HI/*Sst*I fragment of the genomic clone LCFg13ba8a DNA and a 6.8-kb *Bam*HI/*Sst*I fragment of the genomic clone LCFg24ba2b DNA would be candidates for lambda cotton genomic DNAs that encompassed a larger segment of the 5′-flanking region of the genomic *FAD2* gene.

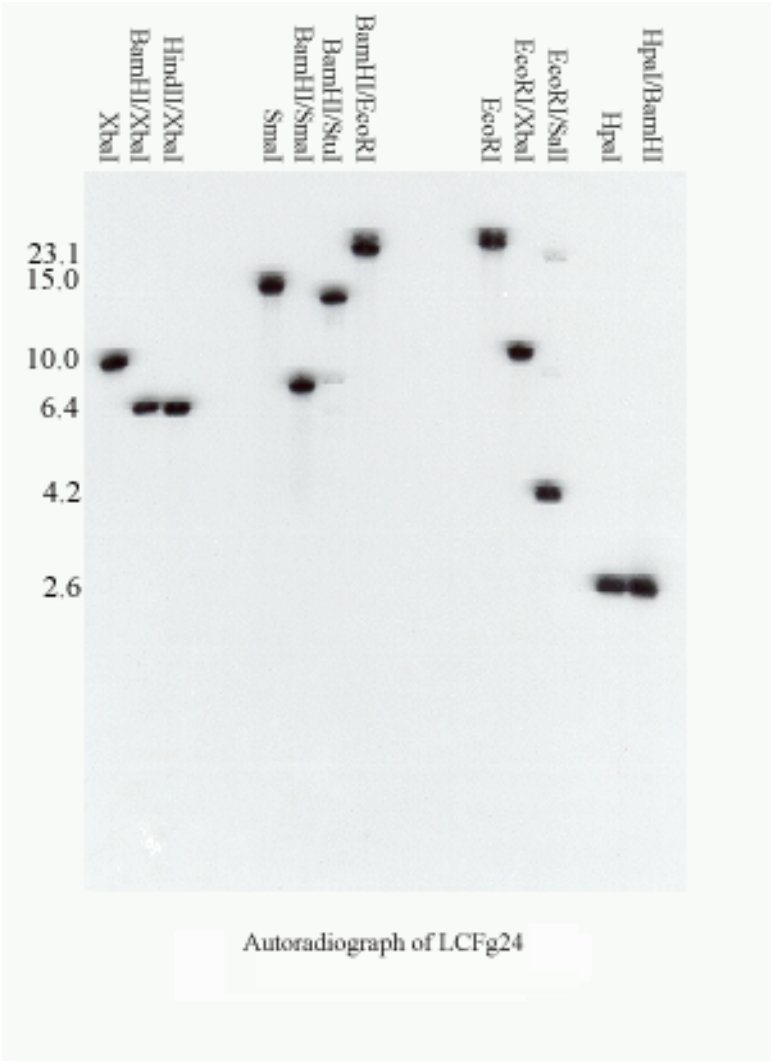
Because of the related restriction fragment patterns of the *Bam*HI/*Sst*I digests of LCFg13ba8a and LCFg24ba2b (Figure 17), these clones probably have the same cotton DNA inserts. Also, from the alkaline blot hybridization (Figure 18), the clone LCFg24ba2b seems to have a longer insert, since the 6.8-kb genomic insert in LCFg24ba2b is larger than the 5.6-kb genomic fragment in LCFg13ba8a, as shown in Figures 17 and 18. The genomic clone LCFg24ba2b was renamed LCFg24 and was selected for physical mapping and subcloning since the 6.8-kb *Hin*dIII fragment was thought to contain the *FAD2* gene 5′-untranslated region, coding region, and 3′-untranslated region. LCFg24 DNA was digested with various restriction endonucleases, as shown in Figures 19 and 21. The two corresponding alkaline blots were hybridized with ^{32}P-labeled 5′-flanking probe. The hybridization profiles are shown in Figure 20 and Figure 22. For example, 2.6-kb *Hpa*I and 2.9-kb *Hin*dIII/*Sal*I fragments were small fragments predicted from the physical map of the clone LCFg55. A comparison of the physical maps of LCFg24 and LCFg55 and their subclones, pCFg24H, pCFg24S, and pCFg55 is shown in Figure 23. LCFg24 has a 12.3-kb cotton DNA fragment harbored in lambda FIXII, whereas LCFg55 has a 13.2-kb cotton DNA harbored in lambda EMBL3. The *FAD2* gene of LCFg24 is located adjacent to the left arm of lambda FIXII while the *FAD2* gene of LCFg55 was nearby the lambda right arm of EMBL3. The LCFg24 has a

Figure 19. Agarose gel electrophoresis of restriction enzyme fragments from the cotton genomic clone LCFg24 DNA on an 0.8% agarose gel. The lambda LCFg24 DNA (200 ng per reaction) was digested with restriction endonucleases. The DNA fragments were stained with ethidium bromide before photography. Each band of the digests of LCFg24 DNA was measured to determine its relative mobility in order to deduce the size of the fragment. The physical map was constructed from analyses of these single and double digestions with a variety of restriction endonucleases. Standard molecular weight markers (1µg) were *Hin*dIII-digested lambda DNA fragments and pGEM DNA markers (Promega). The sizes in kb of these markers are shown at the left.



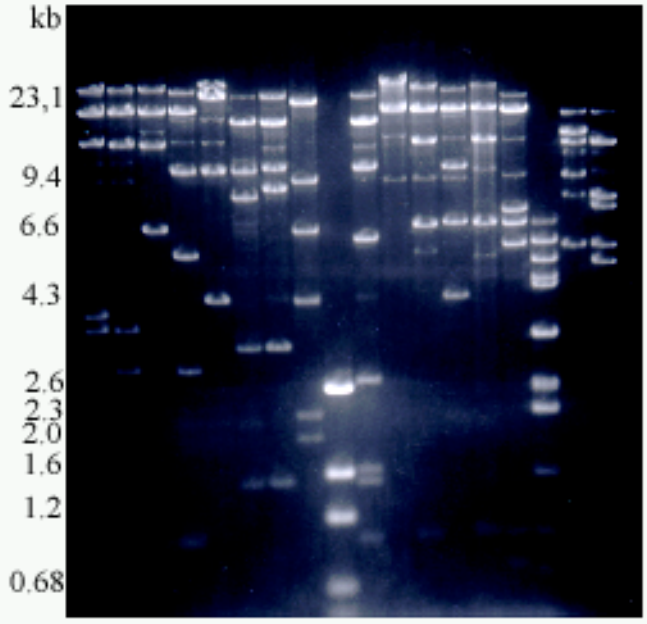
Physical Map of LCFg24 DNA Gel

Figure 20. The alkaline blot hybridization analysis of LCFg24 DNA fragments. LCFg24 DNA was digested with restriction endonucleases (shown in Figure 19) and transferred to a positively-charged nylon membrane. The blot was probed with the ^{32}P-labeled *FAD2* 5′-flanking region probe. The fragments shown in this autoradiogram were the DNA fragments that contained portions of the *FAD2* coding region and 5′-untranslated regions of the *FAD2* gene in LCFg24.



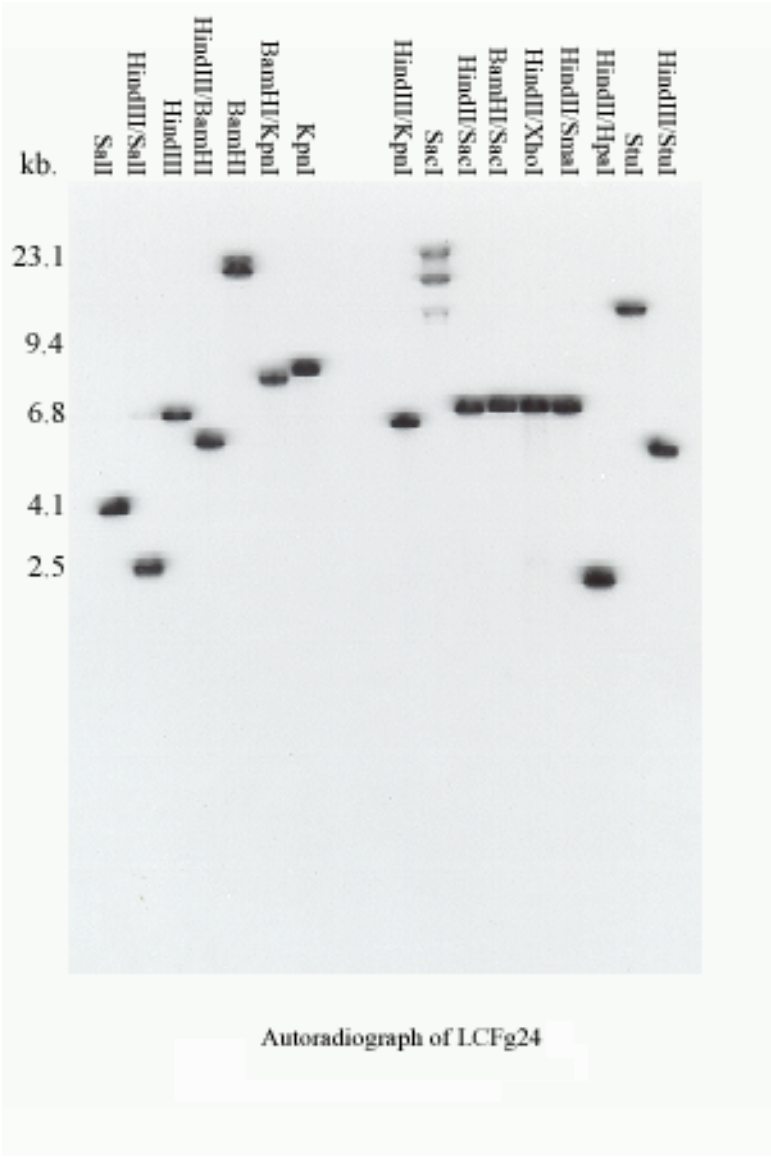
Autoradiograph of LCFg24

Figure 21. Agarose gel electrophoresis of restriction fragments derived from digests of lambda LCFg24 DNA on an 0.8% agarose gel. The single and double digestions of a variety of restriction endonucleases of LCFg24 DNAs (200 ng per reaction) were fractionated. The DNA fragments were stained with ethidium bromide prior to photography. *Hin*dIII-digested Lambda DNA fragments and pGEM DNA markers (1μg) were the standard markers (shown in the middle lanes). The sizes in kb of these markers are shown at the left.



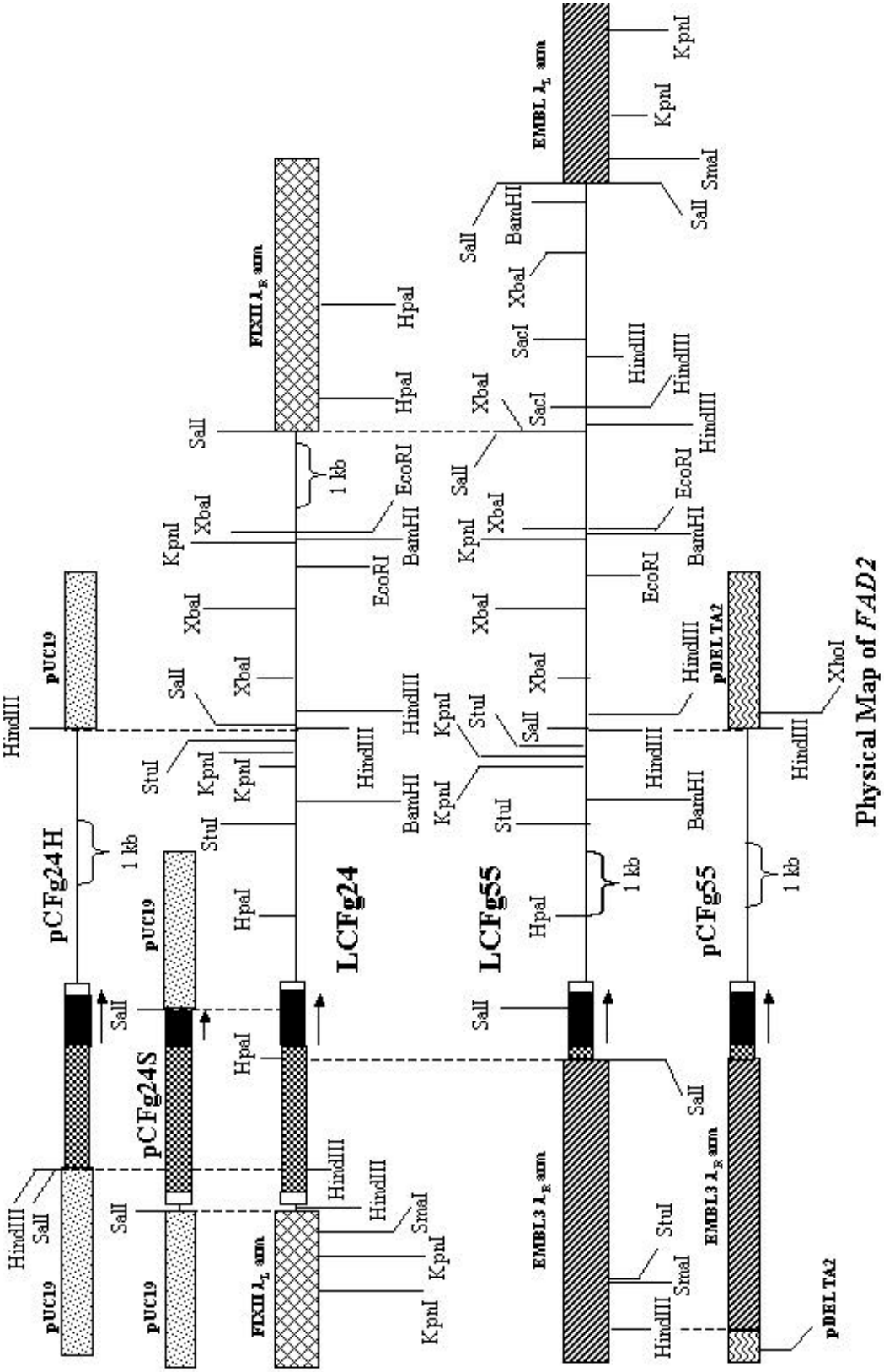
Restriction Mapping of LCFg24

Figure 22. Autoradiogram of the alkaline blot hybridization of an 0.8% agarose gel with restriction enzyme fragments of the cotton genomic clone LCFg24 DNA in Figure 21. The DNA fragments were transferred to a positively-charged nylon membrane, and probed with the ^{32}P-labeled *FAD2* 5′-flanking region probe. The intense hybridization signals shown in each lane represent the DNA fragments that contain the *FAD2* 5′-flanking region and coding region.



Autoradiograph of LCFg24

Figure 23. Physical maps of the cotton genomic clone LCFg24 and the related overlapping genome clone LCFg55. The two panels in the middle represents the map of LCFg24 with 12.3-kb of cotton genomic DNA harbored in lambda FIXII vector (cross-hatched box), and the map of LCFg55 with 13.2-kb of cotton genomic DNA fragment harbored in lambda EMBL3 vector (hatched boxes). The first panel shows the 6.8-kb *Hin*dIII DNA fragment of LCFg24 which was subcloned into the plasmid vector pUC19, and designated pCFg24H. The second panel shows the 4.1-kb *Sal*I DNA fragment of LCFg24 which was subcloned into pUC19, and renamed pCFg24S. The bottom panel is the subclone of LCFg55, named pCFg55, which was subcloned into the cosmid vector pDELTA2.



Physical Map of *FAD2*

longer 5′-untranslated region as demonstrated by the *Sal*I fragment (which extends from the *Sal*I site in both vectors to the *Sal*I in the coding region) since LCFg24 has a 4.2-kb *Sal*I fragment compared to the shorter 1.5-kb *Sal*I fragment of LCFg55. Also, LCFg55 has an extra segment (4.1 kb) of DNA in the 3′-untranslated region that is not found in LCFg24.

As shown in Figure 19, a 6.8-kb *Hin*dIII and a 4.1-kb *Sal*I fragment from LCFg24 hybridize to the ^{32}P-labeled *FAD2* gene-specific, 396-bp 5′-flanking region probe. Therefore, this 6.8-kb *Hin*dIII fragment and the 4.1-kb *Sal*I fragment were subcloned into the *Hin*dIII site and *Sal*I site of the vector of pUC19. Two recombinant plasmids bearing the 6.8-kb *Hin*dIII and 4.1-kb *Sal*I fragments were selected by the blue-white system of selection of the interrupted *lacZ* gene in the pUC19 plasmid. The resulting subclones (named pCFg24H, and pCFg24S) are shown in Figure 23 in the first and second panels of the physical map.

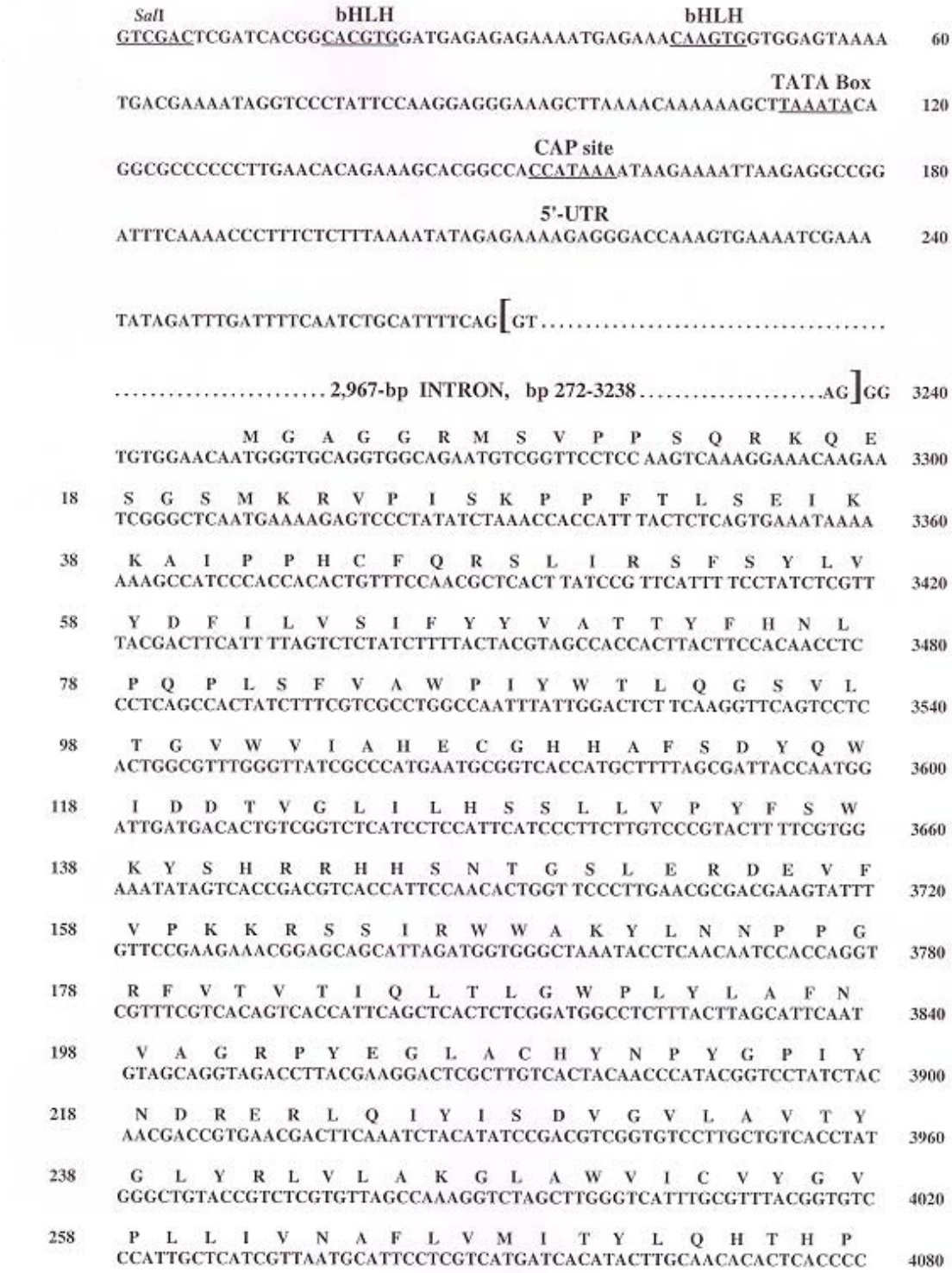
The pCFg24H plasmid DNA was digested with *Hin*dIII to produce the 6.8-kb *Hin*dIII insert, and the presence of the *FAD2* gene was confirmed by alkaline blot hybridization with the ^{32}P-labeled *FAD2* 5′-flanking region probe. The insert in the subclone pCFg24H was used to generate nested random deletions by the EZ::TN Plasmid-Based Deletion Machine method (Epicentre) and sequenced by Dr. John Knesek at Texas Woman's University using the SequiTherm EXCEL II Long-Read DNA Sequencing kit (Epicentre Technologies) with a LI-COR 4800L semi-automated DNA sequencer and by Dr. Irma Pirtle of our laboratory using manual sequencing. The pCFg24S was sequenced on the LI-COR using a primer-based approach with primers

purchased from LI-COR (Lincoln, NE). The DNA sequences were analyzed using DNASIS software (Hitachi).

VII. Characterization of the overlapping cotton genomic clone LCFg24

The sequence analysis and physical mapping confirm that the cotton genomic segments in the clones LCFg24 and LCFg55 both contain the *FAD2* coding regions which are identical within the overlapping regions. The LCFg24 DNA contains a cotton genomic segment (7,914 bp) encompassing the *FAD2* gene, as shown in Figure 24, including the coding region, the 3′-flanking region, and the 5′-flanking region containing the promoter motif and intron. The *FAD2* coding region of 1,155 bp (including the termination codon) is continuous and has no introns. The 5′-flanking region of the *FAD2* gene has a large intron of 2,967 bp, located only 11 nt from the ATG initiation codon (Figure 24). The 3′-untranslated region of the *FAD2* gene is 242 bp, ascertained by comparison with the 3′-flanking region of the cotton *FAD2* partial cDNA clone pSKcF106A described earlier. The presumptive polyadenylation signal could be the hexanucleotide sequence AATCAA, nt 4612-4617 (underlined), which occurs 30 bp upstream from the 3′-polyadenylation site in the *FAD2* mRNA at the underlined A in Figure 24. The *FAD2* gene in LCFg24 has the entire intron (2,967 bp) as compared to the FAD2 gene in LCFg55, which has only 568 bp of the intron. The location and size of the 2,967-bp intron in the 5′-flanking region of the *FAD2* gene was deduced from a comparison of the 5′-end of a cotton *FAD2* cDNA sequence generated by polymerase chain reaction (PCR) amplification of a cotton cDNA library (unpublished results of Mr.

Figure 24. Nucleotide and amino acid sequence of the cotton *FAD2* gene from the clone LCFg24. The numbering on the right refers to nucleotide residues, and the numbering on the left refers to amino acid residues in the presumptive FAD2 polypeptide. The *FAD2* open reading frame has 1,155 bp encoding 384 amino acids. The amino acid sequence of the FAD2 polypeptide is indicated above the coding region in one letter amino acid abbreviations. The intron sequence (demarcated by large brackets) with the 5′-and 3′-splice sites (GT…AG) at nt 272 and 3,238, respectively, is not shown. The CAP site or the 5′ end of the mature mRNA is located at nt 153. The 5′-flanking untranslated region (5′-UTR) of the mRNA is 130 nt. A possible TATA box and two basic region helix-loop-helix residues (bHLH) or E box motifs are underlined upstream from the CAP site at nt 113, 135, and 109, respectively. The putative polyadenylation signal (AATCAA) is underlined at nt 4,612-4,617. The 3-untranslated region between the poly A site at nt 4,247 and nt 7,860 has been omitted for brevity. This cotton *FAD2* gene sequence has been assigned GenBank Number AF331163.


SalI
bHLH
bHLH
GTCGACTCGATCACGGCACGTGGATGAGAGAGAAAATGAGAAACAAGTGGTGGAGTAAAA 60
TATA Box
TGACGAAAATAGGTCCCTATTCCAAGGAGGGAAAGCTTAAAACAAAAAAGCTTAAATACA 120
CAP site
GGCGCCCCCCTTGAACACAGAAAGCACGGCCACCATAAAATAAGAAAATTAAGAGGCCGG 180
5'-UTR
ATTTCAAAACCCTTTCTCTTTAAAATATAGAGAAAAGAGGGACCAAAGTGAAAATCGAAA 240
TATAGATTTGATTTTCAATCTGCATTTTCAG[GT.....
.....2,967-bp INTRON, bp 272-3238.....AG]GG 3240
M G A G G R M S V P P S Q R K Q E
TGTGGAACAATGGGTGCAGGTGGCAGAATGTCGGTTCCTCC AAGTCAAAGGAAACAAGAA 3300
18 S G S M K R V P I S K P P F T L S E I K
TCGGGCTCAATGAAAAGAGTCCCTATATCTAAACCACCATT TACTCTCAGTGAAATAAAA 3360
38 K A I P P H C F Q R S L I R S F S Y L V
AAAGCCATCCCACCACACTGTTTCCAACGCTCACT TATCCG TTCATTT TCCTATCTCGTT 3420
58 Y D F I L V S I F Y Y V A T T Y F H N L
TACGACTTCATT TTAGTCTCTATCTTTTACTACGTAGCCACCACTTACTTCCACAACCTC 3480
78 P Q P L S F V A W P I Y W T L Q G S V L
CCTCAGCCACTATCTTTCGTCGCCTGGCCAATTTATTGGACTCT TCAAGGTTCAGTCCTC 3540
98 T G V W V I A H E C G H H A F S D Y Q W
ACTGGCGTTTGGGTTATCGCCCATGAATGCGGTCACCATGCTTTTAGCGATTACCAATGG 3600
118 I D D T V G L I L H S S L L V P Y F S W
ATTGATGACACTGTCGGTCTCATCCTCCATTCATCCCTTCTTGTCCCGTACTT TTCGTGG 3660
138 K Y S H R R H H S N T G S L E R D E V F
AAATATAGTCACCGACGTCACCATTCCAACACTGGT TCCCTTGAACGCGACGAAGTATTT 3720
158 V P K K R S S I R W W A K Y L N N P P G
GTTCCGAAGAAACGGAGCAGCATTAGATGGTGGGCTAAATACCTCAACAATCCACCAGGT 3780
178 R F V T V T I Q L T L G W P L Y L A F N
CGTTTCGTCACAGTCACCATTCAGCTCACTCTCGGATGGCCTCTTTACTTAGCATTCAAT 3840
198 V A G R P Y E G L A C H Y N P Y G P I Y
GTAGCAGGTAGACCTTACGAAGGACTCGCTTGTCACTACAACCCATACGGTCCTATCTAC 3900
218 N D R E R L Q I Y I S D V G V L A V T Y
AACGACCGTGAACGACTTCAAATCTACATATCCGACGTCGGTGTCCTTGCTGTCACCTAT 3960
238 G L Y R L V L A K G L A W V I C V Y G V
GGGCTGTACCGTCTCGTGTTAGCCAAAGGTCTAGCTTGGGTCATTTGCGTTTACGGTGTC 4020
258 P L L I V N A F L V M I T Y L Q H T H P
CCATTGCTCATCGTTAATGCATTCCTCGTCATGATCACATACTTGCAACACACTCACCCC 4080


278 A L P H Y D S S E W D W L R G A L A T V
GCAT TACCACACTACGACTCATCCGAATGGGACTGGTTACGTGGAGCCCTCGCGACGGTC 4140
298 D R D Y G I L N K V F H N I T D T H V A
GACCGAGATTATGGGATATTAAACAAGGTTTTCCATAACATAACTGATACTCATGTCGCT 4200
318 H H L F S T M P H Y H A M E A T K A I K
CATCATTTGTTTTCGACGATGCCGCATTACCACGCAATGGAAGCAACTAAGGCAATAAAA 4260
338 P I L G E Y Y S F D G T P V Y K A I F R
CCAATATTGGGAGAGTATTATTCATTTGATGGTACACCAGTTTATAAAGCGATATTTAGA 4320
358 E A K E C I Y V E P D E G E Q S S K G V
GAGGCAAAGGAGTGTATTTACGTTGAACCAGACGAAGGTGAGCAGAGCAGCAAAGGTGTA 4380
378 F W F R N K I TER
TTTTGGTTTAGAAATAAGATCTAACTTTGCCGATAGCGTTGCGGTTGCCGATGGTGATGC 4440
CTTTAGGAATGTGTTAAATTTGT TACATTAT TGTTAAGGATTTGGGGTTTTGGCGTTTGG 4500
3'-UTR
GTTACATCAATTTCAGATGCTTTCGAATTTGGACTTTGTATGGT TCTCATCGACTTTGTT 4560
Poly A Signal
GATCCCTGCAAAATTGGT TCGAGCTTTCAACTATCAAGTAGTTTTTTTTTTAATCAAATT 4620
Poly A site
TATTATTGGTGCCGAGT TATAAAAAA..... 4680
.....
GACAACCCGGTCTTGGAAACAAGGATCAGGGGATTCAAGCAAGGCTTGAAGCTT 7914
HindIII

Mongkol Nampaisansuk, Dr. Kent Chapman, and Dr. Irma Pirtle of this research group) and the *FAD2* gene sequence in LCFg24. The presumptive transcription initiation site corresponds to the 5′-end or CAP site of the mature *FAD2* mRNA (CCATAAA) at nucleotide (nt) 153 in Figure 24. Since the 2,967-bp intron is spliced out to generate the mature *FAD2* mRNA, the 5′-flanking untranslated region (5′-UTR) of the mature mRNA would be 130 nucleotides. The 5′-flanking intron occurs only 11 bp upstream from the ATG initiation codon. The *Arabidopsis FAD2* gene (Okuley et al., 1994) also has a rather large, 1,100-bp intron in its 5′-flanking region, only four basepairs from its ATG initiation codon. The presence of large introns in the 5′-flanking regions of *FAD2* genes could be important in the transcriptional regulation of expression of these genes.

The 5′-flanking region of the *FAD2* gene has several potential promoter/enhancer elements that could function as positive regulatory elements in gene expression. A TATA basal promoter element occurs 41 bp upstream from the putative cap site. Two basic region helix-loop-helix (bHLH) or E box motifs with the consensus sequence CANNTG occur at 109 bp and 135 bp upstream from the potential cap site. The E box motif has been shown to be a seed-specific regulatory element in the French bean β-phaseolin gene (Kawagoe et al., 1994).

VIII. Genomic organization of cotton *FAD2* genes

The determination of the number of *FAD2* genes in a number of plant genomic DNAs has been performed by genomic blot hybridization. Heppard et al. (1994) did genomic blot analyses of soybean *FAD2-1* and *FAD2-2* desaturase genes and found at least two copies of these *FAD2* genes in the soybean genome. Soybean genomic DNA

was digested with *Bam*HI and *Hin*dIII and subjected to alkaline blot hybridization with soybean *FAD2-1* and *FAD2-2* gene-specific probes. However, the *Arabidopsis FAD2* gene is a single-copy gene (Okuley et al., 1994). Plant desaturase genes seem to be single copy, except for the omega-3-desaturase gene or *FAD3* gene (Gibson et al., 1994). Since the soybean and cotton genomes are allotetraploid (Heppard et al., 1996), it is possible that four (or more) copies of the *FAD2* genes could be found in these two plant genomes and that the two pairs of genes might be nonallelic. To determine the copy number of the *FAD2* genes in cotton genomic DNA, the coding region of the cotton *FAD2* gene was used as a probe for a genomic blot.

The 2.6-kb *Hpa*I, 2.9-kb *Hin*dIII/*Sal*I, and 6.8-kb *Hin*dIII fragments indicated in the physical map of the genomic clone LCFg24 in Figure 23 are DNA fragments encompassing the *FAD2* coding region. As shown in Figure 25, digests of cotton genomic DNAs and LCFg24 DNA with *Hpa*I, *Hin*dIII/*Sst*I, and *Hin*dIII were done, fractionated on an agarose gel, and alkaline-blotted to a positively-charged nylon membrane. Restriction fragments on the alkaline blot of actual cotton genomic DNA and the cloned lambda LCFg24 DNA were hybridized overnight at 60˚C under high stringency conditions to a ^{32}P-labeled probe generated by random priming using a 1.2-kb *Eco*RI/*Sac*I DNA fragment corresponding to the FAD2 coding region from the plasmid construct pYES2/FAD2 (shown in Figure 27).

The LCFg24 DNA fragments digested with *Hpa*I, *Hin*dIII/*Sal*I, and *Hin*dIII shown in Figure 25 correspond to the DNA fragments shown in the physical maps of LCFg24S, and pCFg24H depicted in Figure 23. Panel B of Figure 25 shows the

smeared fragments of cotton genomic DNA digested with *Hpa*I, *Hin*dIII-*Sal*I, and *Hin*dIII. The ^{32}P-labeled probe generated from the 1.2-kb cotton DNA *Eco*RI/*Sac*I fragment from the *FAD2* coding region hybridized to genomic fragments in the alkaline blot (Figure 26). Panel A of Figure 26 shows the autoradiogram of the alkaline blot hybridization of the cloned LCFg24 DNA fragments from Panel A of Figure 25. The 2.6-kb *Hpa*I fragment that hybridized to the *FAD2* probe encompassed the entire *FAD2* coding region. This means that *Hpa*I cleaved the *FAD2* gene in the intron of the 5′-flanking region and in the 3′-untranslated region.

*Hin*dIII digestion generated a 6.8-kb fragment corresponding to the LCFg24 DNA fragment that was cleaved in the intron and in the middle of the cotton DNA insert. The double digestion of *Hin*dIII and *Sal*I generated an intensely hybridizing 2.9-kb *Hin*dIII/*Sal*I fragment that contained most of the coding region, and a faintly hybridizing 3.9-kb *Hin*dIII/*Sal*I fragment encompassing a minor portion of 3′-coding region. In Figure 26 (panel B), two distinct bands of cotton genomic DNA occur after digestion with *Hin*dIII (6.8 kb and 7.8 kb) and with double digestion with *Hin*dIII and *Sal*I (2.9 kb and 7.8 kb). The two hybridizing fragments (6.8-kb *Hin*dIII and 2.9-kb *Hin*dIII/*Sal*I) directly correspond to the 6.8-kb *Hin*dIII and 2.9-kb *Hin*dIII/*Sal*I fragments from the cloned LCFg24 DNA. The 7.8-kb fragment with *Hin*dIII-*Sal*I would correspond to the 7.8-kb fragment with *Hin*dIII in Panel B. The blot of the double digestion of cotton genomic DNA with *Hin*dIII and *Sal*I supports the fact that the 6.8-kb *Hin*dIII fragment from LCFg24 had disappeared, and that the 2.9-kb fragment would correspond to the 2.9-kb *Hin*dIII/*Sal*I fragment from LCFg24 (the middle lane of Panel

Figure 25. Agarose gel electrophoresis of cotton genomic DNA restriction fragments. The cloned LCFg24 DNA (2 µg per reaction) and cotton genomic DNA (15 µg per reaction) were digested with *Hin*dIII, *Hin*dIII/*Sal*I, and *Hpa*I, and the fragments resolved on an 0.8% agarose gel. The DNA fragments were stained with ethidium bromide prior to photography. Panel A shows the restriction fragments from LCFg24 DNA. Panel B shows the cotton genomic DNA restriction fragments.

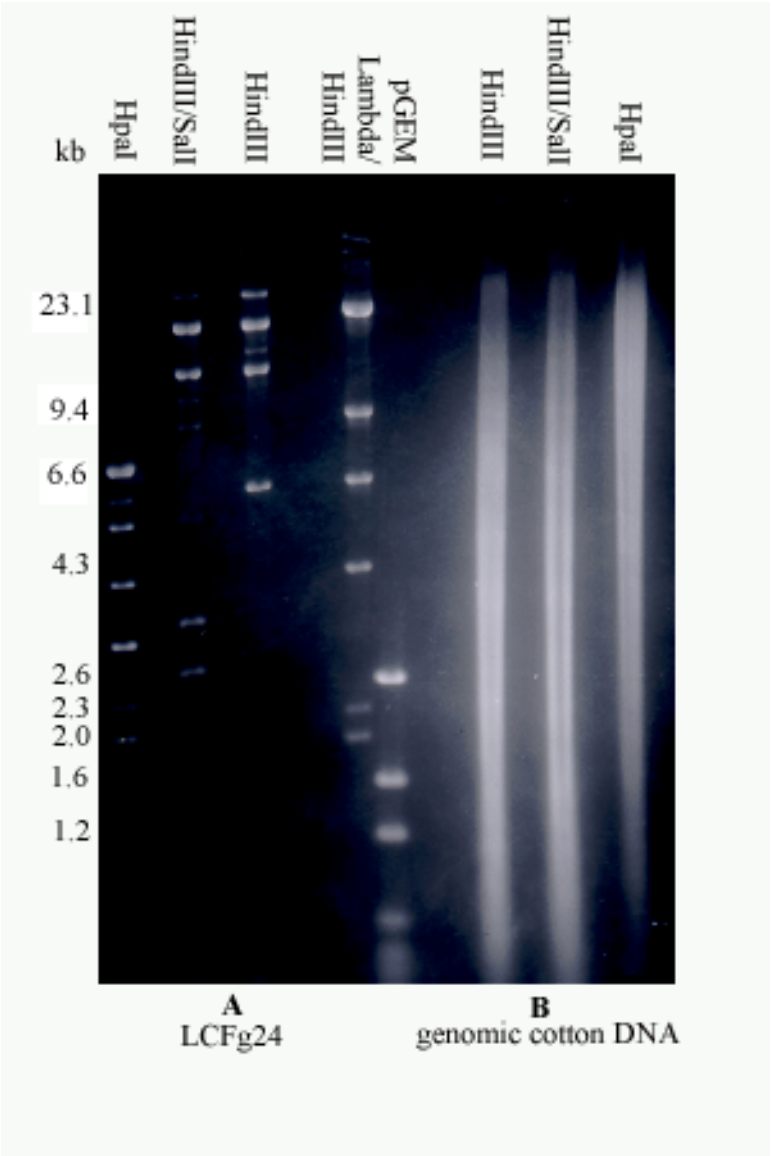
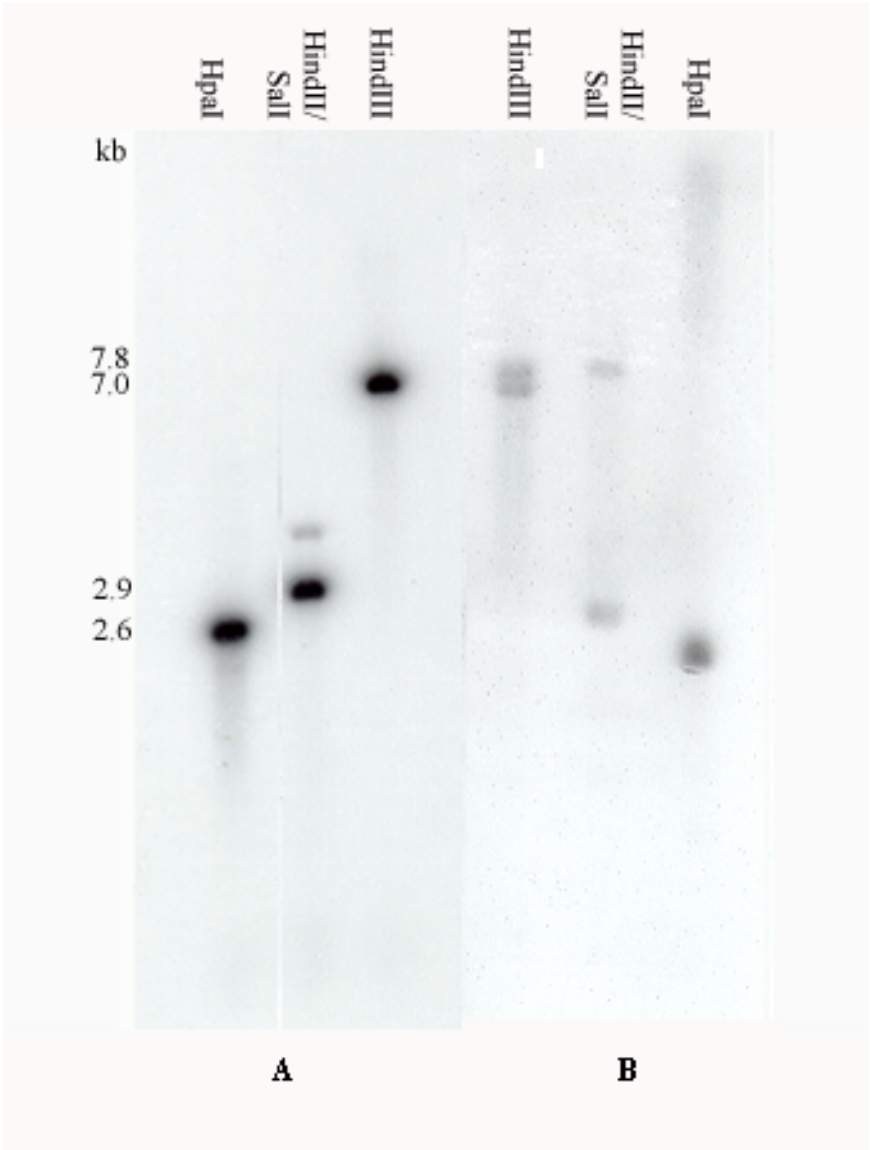

kb
HpaI
HindIII/SalI
HindIII
pGEM Lambda/ HindIII
HindIII
HindIII/SalI
HpaI
23.1
9.4
6.6
4.3
2.6
2.3
2.0
1.6
1.2
A
LCFg24
B
genomic cotton DNA

Figure 26. Autoradiogram of alkaline blot hybridization of cotton genomic DNA. As shown in Figure 25, the LCFg24 and cotton genomic DNAs were digested with restriction endonucleases (*Hpa*I, *Hin*dIII/*Sal*I, *Hin*dIII), fractionated on an 0.8% agarose gel, and transferred to a positively-charged nylon membrane. The DNA fragments were hybridized to ^{32}P-labeled probe derived from the 1.2-kb *Eco*RI/*Sac*I fragment containing the *FAD2* coding region of pYES2/FAD2. Panel A shows the hybridizing *Hpa*I, *Hin*dIII/*Sal*I, *and Hin*dIII fragments of LCFg24 DNA. Panel B shows the unique hybridizing *Hpa*I, *Hin*dIII/*Sal*I, and *Hin*dIII fragments encompassing the *FAD2* gene in cotton genomic DNA.


HpaI
SalI
HindII/
HindIII
HindIII
SalI
HindII/
HpaI
kb
7.8
7.0
2.9
2.6
A
B

B in Figure 26). The 2.6-kb *Hpa*I hybridizing fragment would correspond to the 2.6-kb *Hpa*I fragment of LCFg24 DNA encompassing the *FAD2* coding region. Since the 6.8-kb *Hin*dIII, 2.9-kb *Hin*dIII/*Sal*I, and 2.6-kb *Hpa*I fragments from the cloned LCFg24 DNA are identical to the sizes of the cotton genomic DNA fragments, the cloned cotton insert in the genomic clone LCFg24 must represent actual cotton genomic DNA without any cloning artifacts. The 7.8-kb *Hin*dIII and 7.8-kb *Hin*dIII/*Sal*I fragments must be derived from a second *FAD2* gene in the cotton genome. Consequently, this indicates that there are at least two (or more) *FAD2* genes in the allotetraploid cotton genome, which would be similar to the arrangement of the soybean *FAD2* genes in its allotetraploid genome.

IX. Expression of the cotton *FAD2* coding region in yeast cells

A number of plant microsomal *FAD2* and *FAD3* cDNAs and genes have been isolated and characterized, but the role and biochemical functions of the expressed polypeptide products is still largely uncharacterized. The role of these enzymes can be studied in *vivo* in a yeast expression system. The functional expression of the cotton *FAD2* gene should be possible in yeast cells since they produce only the monounsaturated fatty acid palmitoleic (16:1) and oleic (C18:1) acids with the *OLE1* gene (Mitchell and Martin, 1995). The synthesis of the polyunsaturated fatty acid species linoleic (C18:2) and linolenic (C18:3) acids does not occur in yeast cells since they do not have the FAD2 or FAD3 enzymes found in plants. For example, the *Arabidopsis FAD2* gene was expressed in yeast using the *GAL1* promoter of *Saccharomyces cerevisiae* (Kajiwara et al., 1996; Covello and Reed, 1996). Yeast cells

form the hexadecadienoyl (C16:2) and linoleoyl (C18:2) residues in the yeast membrane phospholipids when the *Arabidopsis FAD2* coding region is induced by galactose. The expression of the cotton *FAD2* gene described in this dissertation was analyzed in transformed yeast cells as previously done for *Arabidopsis FAD2* gene expression in yeast cells (Covello and Reed, 1996; Kajiwara et al., 1996).

The cotton *FAD2* coding sequence (for 384 amino acids plus start and stop codons) was amplified by PCR using primers as described in Chapter II. The PCR product was inserted into the yeast-bacterial shuttle expression vector pYES2 (Invitrogen) just downstream of the galactose-inducible *GAL1* promoter, as shown in Figure 27, and the construct was designated pYES2/FAD2. The pYES2/FAD2 construct was transformed into *E. coli* DH5α. The sequence of the *FAD2* coding sequence inserted in pYES2/FAD2 was confirmed by manual sequencing by Dr. Irma Pirtle of our laboratory.

The plasmid pYES2/FAD2 was transformed into the *S. cerevisiae* strain INVSc1 (Invitrogen) by the lithium acetate method (Gietz, 1992) and electroporation (Becker, 1991). The expression of this recombinant plasmid was induced by addition of galactose to up-regulate the yeast *GAL1* promoter. The pYES2 (as control) and pYES2/FAD2 transformants were grown at 30°C overnight in a medium containing raffinose as the sole carbon source. Then, the total lipids were extracted and analyzed by gas chromatography by our collaborator, Dr. Kent Chapman, of our department (Chapman and Trelease, 1991). Chromatograms of GC profiles of the fatty acid methyl esters are shown in Figure 28. Linoleic acid (C18:2) and palmitoleic acid (C16:2) were

not detected in the control yeast cells with pYES2 when no galactose was present. The expression of pYES2/FAD2 with the yeast *GAL1* promoter was induced by addition of 2% galactose. The pYES2 and pYES2/FAD2 transformants were grown at 30°C for 22 hours (three generations). When the fatty acid methyl esters were extracted and analyzed by GC analysis, the pYES2/FAD2 yeast transformants had a significant accumulation of C18:2 not normally present in wild-type yeast cells and a small amount of C16:2 as shown in Figure 29. In addition, the oleic acid (C18:1) peak in the transformed yeast cells containing the cotton *FAD2* gene (Figure 29) was significantly smaller than the C18:1 peak in the control yeast cells of Figure 28. Moreover, the C18:2 and C16:2 peaks were detected in two separate experiments, in the SD medium containing 2% galactose with added exogenous oleic acid in the presence of a low concentration of the detergent tergitol (not shown). Thus, oleic acid (C18:1) was clearly converted to linoleic acid (C18:2) in the yeast cells containing the plasmid construct with the cotton *FAD2* gene. Thus, this experiment indicates that the *FAD2* gene is truly expressed in the yeast *S. cerevisiae*, and that the cotton *FAD2* gene described in this dissertation is indeed is a functional gene.

Figure 27. The structure of the plasmid construct designated pYES2/FAD2 transformed into *Saccharomyces cerevisiae* strain INVSc1. This plasmid construct contains the cotton *FAD2* gene and was constructed from the pYES2 (Invitrogen) plasmid. The *FAD2* gene was inserted into the *Sac*I-*Eco*RI sites of the polylinker region located at the 3′ end of the *GAL* promoter. The pYES2 vector also contains the CYC1 transcription terminator, ampicillin resistance gene marker, yeast URA3 gene marker, the bacterial pMB1 (pUC-derived) origin of replication (pMBI ori), f1 origin of replication (fi ori), and yeast 2 micron origin of replication (2u ORI).

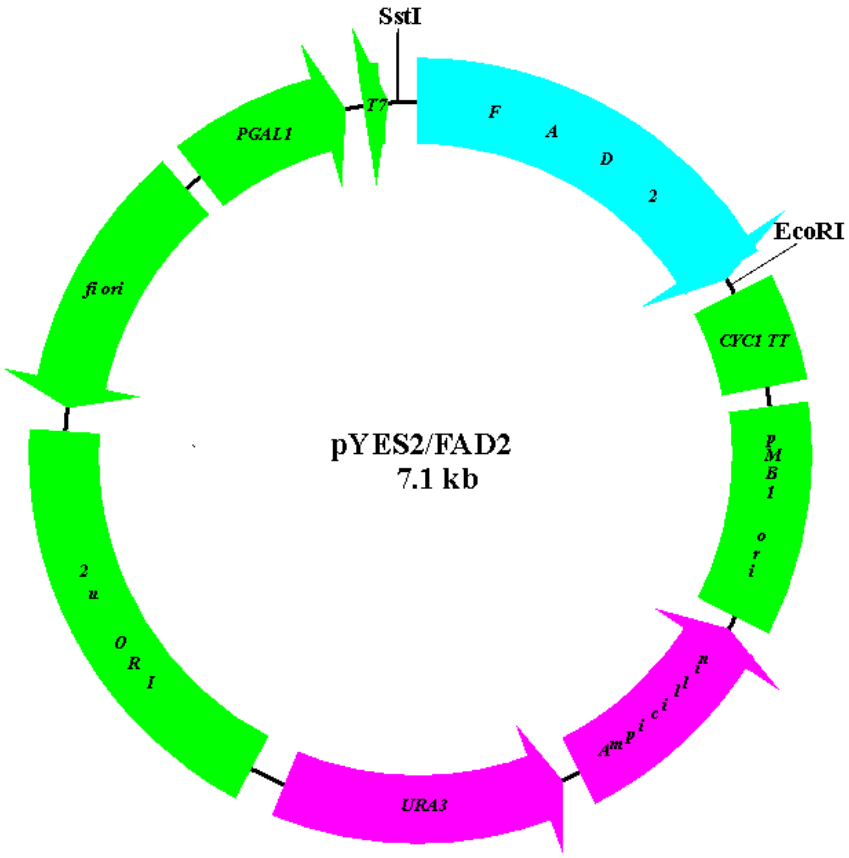

SstI
T7
PGAL1
F
A
D
2
EcoRI
CYC1 TT
pMB1 ori
URA3
Ampicillin
2 micron ORI
f1 ori
pYES2/FAD2
7.1 kb

Figure 28. Gas chromatographic analysis of the total fatty acid methyl esters extracted from the control *S. cerevisiae* transformed with the pYES2 plasmid. The gas chromatography (GC) was done using flame ionization detection. The *S. cerevisiae* pYES2 transformants were grown in SD dropout uracil medium without galactose at 30ºC for 3 generations. The fatty acid peaks were identified as C16:0, C16:1, C17:0 (internal standard), C18:0, and C18:1.

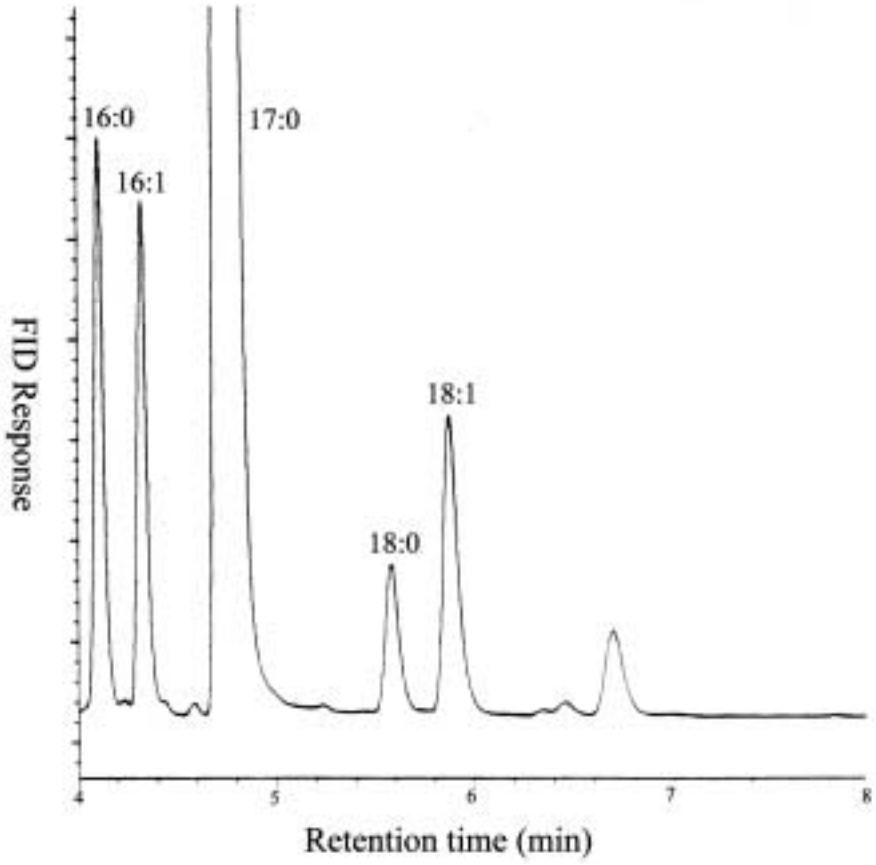
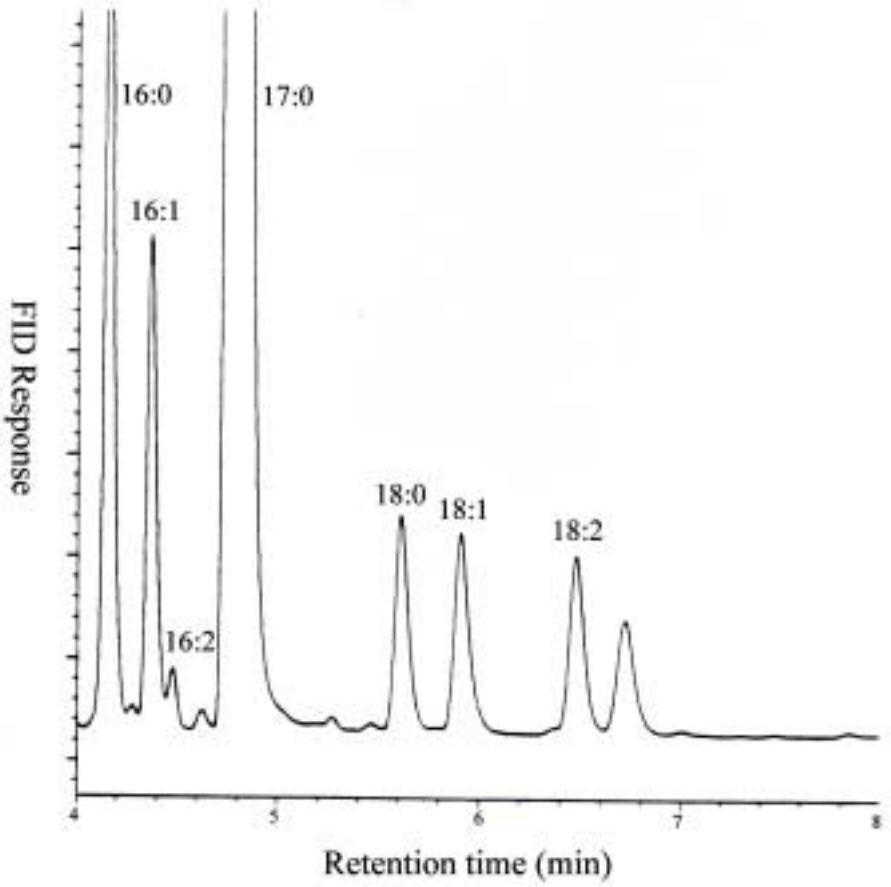

16:0
16:1
17:0
18:0
18:1
FID Response
4
5
6
7
8
Retention time (min)

Figure 29. Gas chromatographic analysis of the total fatty acid methyl esters extracted from *S. cerevisiae* transformed with the plasmid construct pYES2/FAD2. The gas chromatography (GC) was done using flame ionization detection. The *S. cerevisiae* pYES2/FAD2 transformant was grown in SD dropout uracil with 0.2% galactose at 30˚C for 3 generations. The peaks are C16:0, C16:1, C16:2, C17:0 (internal standard), C18:0, C18:1, and C18:2.


16:0
17:0
16:1
18:0
18:1
18:2
16:2
FID Response
4
5
6
7
8
Retention time (min)

CHAPTER IV

DISCUSSION

I. Molecular cloning of the gene encoding a cotton ω-6 fatty acid desaturase (FAD2)

Two cotton genomic clones encompassing the same fatty acid desaturase (FAD2) gene and a partial-length cotton *FAD2* cDNA clone were successfully isolated. They were identified from their DNA sequence identities with the cDNAs/genes for other plant fatty acid desaturases, especially the *Arabidopsis FAD2* gene. First, we isolated a *FAD2* cDNA clone from a cotton (*Gossypium hirsutum* cv. Deltapine 62) cDNA library (Ni and Trelease, 1991) using a heterologous 0.5-kb *Sal*I fragment isolated from the *Arabidopsis FAD2* cDNA clone designated *FAD 2-43* (Newman et al., 1994; Okuley et al., 1994). In addition, a cotton genomic library harbored in the lambda vector EMBL3 (Grula et al., 1995) was screened with the *Arabidopsis FAD2* probe. One genomic clone designated LCFg55 was found to have a 13.2-kb cotton genomic fragment encompassing a putative *FAD2* gene. A 10.0-kb *Hin*dIII fragment of LCFg55 with 4.2-kb of EMBL3 vector DNA and 5.8-kb of cotton genomic DNA was subcloned into the cosmid vector pDELTA and designated pCFg55. The subclone pCFg55 was found to have the complete FAD2 protein-coding region (1,155 bp), 568 bp of 5′- untranslated region, and 242 bp of 3′-untranslated region. However, the clone LCFg55 contains an incomplete *FAD2* gene since the genomic insert in this clone was cleaved in the 5′-flanking region of the *FAD2* gene prior to its ligation into the lambda vector EMBL3.

Thus, it became necessary to isolate an overlapping genomic clone containing the *FAD2* gene from a second cotton genomic library.

For screening a second cotton genomic library harbored in the lambda FIXII vector (gift of Dr. Thea Wilkins of the University of California at Davis), a *FAD2* gene-specific 5′-flanking region fragment of 396 bp was generated by PCR amplification and subcloned into the vector pGEM7Zf(+) to prepare ^{32}P-labeled hybridization probe by random priming. A second lambda clone designated LCFg24 that strongly hybridized with the *FAD2* 5′-flanking region probe was isolated and characterized by physical mapping and DNA sequence analysis. The clone LCFg24 has a 12.3-kb cotton genomic DNA fragment harbored in lambda FIXII. Since the two genomic fragments in LCFg24 and LCFg55 overlap by 9.6 kb (shown in the physical maps in Figures 7 and 23), an overall 16.5-kb region of cotton genomic DNA is harbored in the two clones. The 7,914-bp genomic DNA sequence in the clone LCFg24 must represent actual cotton genomic DNA, since the 6.8-kb *Hin*dIII, 2.9-kb *Hin*dIII/*Sal*I, and 2.6-kb *Hpa*I fragments from cloned LCFg24 DNA correspond to identically-sized cotton genomic fragments (shown in Figure 26).

The cotton *FAD2* gene encodes a putative FAD2 polypeptide of 384 amino acid residues. By comparison of the alignments of the deduced amino acid sequences among the plant ω–6 desaturases, the cotton FAD2 protein has the highest amino acid identity (85%) with the cotton FAD2-2 polypeptide (Liu et. al, 1999) and also a high identity (75%) with the cotton FAD2-1 polypeptide (Liu et. al., 1997). Furthermore, the FAD2 polypeptide in this work has high identities (72-85%) with the amino acid sequences of

other plant FAD2 polypeptides (Figure 13). Thus, the cotton *FAD2* gene has definitely been shown to be a new member of the plant *FAD2* gene family. It has been shown that only a single *FAD2* gene in the *Arabidopsis* genome is expressed in both vegetative tissues and developmental seeds (Okuley et al., 1994). In soybean, at least two *FAD2* genes are expressed. The soybean *FAD2-1* gene is expressed in developing seeds, but the *FAD2-2* is expressed in both vegetative tissues and in developing seeds (Heppard et al., 1996). The cotton *FAD2* gene described in this dissertation may be expressed in both vegetative tissues and in developing embryos because of its high identity to the cotton *FAD2-2* cDNA isolated by Liu et al. (1999), which is apparently expressed in both vegetative tissues and developmental seeds (Liu et al., 1999). The regulation of expression of the cotton *FAD2* gene in this dissertation is under investigation in our laboratory at this time by Ms. Stacy Park.

II. Analysis of the cotton *FAD2* gene

The nucleotide sequence encompassing the cotton *FAD2* gene in the genomic clones LCFg24 and LCFg55 is 7,914-bp in length (GenBank Number AF331163). The FAD2 open reading frame has 1,155 bp, including the initiation (ATG) and stop (TAA) codons, encoding a polypeptide of 384 amino acids. The *FAD2* coding region has no intervening sequence or intron. The 5′-flanking region has a 2,967-bp intron located only 11 bp upstream from the ATG initiation codon. This *FAD2* gene has a 3′-untranslated region (3′-UTR) of 242 bp downstream from the UAA termination codon. The polyadenylation signal occurs 30 bp upstream from the 3′-polyadenylation site, shown in Figure 24.

The presumptive promoter/enhancer elements of the *FAD2* gene located in the 5′-flanking region may be necessary for transcriptional regulation of gene expression. The basal promoter element (TATA box) is found 41 nucleotides upstream of the putative CAP site. Prospective seed-specific regulatory elements (E box motifs) are found 109 bp and 135 bp upstream from the potential CAP site. These E box motifs are referred to as the basic region helix-loop-helix (bHLH) for seed-specific gene expression (Kawagoe et al., 1994).

The homology between the DNA sequence of the cotton *FAD2* gene and several other fatty acid desaturase cDNA or gene sequences is indicated in Table I. As previously mentioned, the cotton *FAD2* gene in this research has high identity to the cotton *FAD2-1* and *FAD2-2* genes (62% and 69%, respectively). The cotton *FAD2* gene is highly similar to the *FAD2 Arabidopsis* gene (64% identity) and the *Glycine max FAD2-1* gene (60% identity). Thus, a high degree of sequence similarity is found among the plant ω–6 desaturase (*FAD2*) genes. However, the cotton *FAD2* gene sequence has only low to moderate identities (41%-48%) with other fatty acid desaturase genes, such as the *Arabidopsis FAD3, FAD6, FAD7,* and *FAD8* genes. The cotton *FAD2* DNA sequence does not have high similarities with other classes of fatty acid desaturases, such as the Δ9-fatty acid desaturase (*OLE1*) gene of *S. cerevisiae* (Stukey, 1990), and the Δ6-fatty acid desaturase (*DESD*) of *Synechocystis* sp. (Reddy, 1993). This comparison shows the divergence of fatty acid desaturase DNA sequences among plants, cyanobacteria, and yeast (Okuley et al., 1994; Tocher et al., 1998; Shanklin, 1998).

Table I. Homology between the DNA sequence of the cotton *FAD2* gene and other fatty acid desaturase cDNA /gene sequences

Gene/cDNA Sequence	name	identities	location
Arabidopsis FAD2	Δ12- desaturase	64%	ER
Arabidopsis FAD3	Δ15- desaturase	41%	ER
Arabidopsis FAD6	Δ12- desaturase	48%	plastidial
Arabidopsis FAD7	Δ15-desaturase	39%	plastidial
Arabidopsis FAD8	Δ12-desaturase	40%	plastidial
Brassica napus FAD6	ω-3 desaturase	62%	ER
Boraco officinalis	Δ6-desaturase	48%	ER, seed
Glycine max FAD2-1	ω-6 desaturase	60%	ER
Glycine max FAD2-2	ω-6 desaturase	57%	ER
FAD2-1	ω -6 desaturase	62%	ER
Gossypium hirsutum FAD2-2	ω -6 desaturase	69%	ER
Saccharomyces cerevisiae OLE1	Δ9-desaturase	41%	ER
Synechocistis DESA	Δ12-desaturase	53%	thylakiod membrane
Synechocistis DESD	ω-6 desaturase	46%	thylakoid membrane

In the case of the *Arabidopsis FAD7* gene (Iba et al., 1993) and *FAD8* gene (Gibson et al., 1994), seven introns are located in the protein-coding regions, ranging from 79 to 301 bp in size for the *FAD7* gene and from 84 to 113 bp in size for the *FAD8* gene. In contrast, the cotton *FAD2* gene has one large intron of 2,967 bp located in the 5′- untranslated region, just upstream of its open reading frame.

The *Arabidopsis* FAD7 polypeptide has plastid transit sequences (Tocher et al., 1998; Iba et al., 1993), whereas the cotton FAD2 protein lacks this signal. It has been suggested that not only the primary structures of the desaturase enzymes, but also the targeting signals for sorting the proteins into either the chloroplast or ER, has implications for the divergence of fatty acid desaturase amino acid sequences. The cotton FAD2 protein may need the lysine- or arginine-rich carboxy-terminal motifs which represent the retention signal for integral membrane proteins in the ER membrane (Teasdale et al., 1996). The membrane-bound desaturase superfamily in plants, cyanobacteria, and yeasts, has homology in the three histidine-rich motifs which serve as the active site-diiron cluster for desaturation (Tocher et al., 1998; Shanklin et al., 1998).

A cytochrome b_5 domain occurs at the C-terminal domain of some desaturase enzymes (Stukey et al., 1990; Itoh et al., 1998; Tocher et al., 1998). For example, the *Saccharomyces cerevisiae* OLE1 protein has a cytochrome b_5 domain. The orientation of the cytochrome b_5 domain to the fatty acid desaturase catalytic center could accelerate electron transfer and eliminate the need for diffusion of the NADH cytochrome b_5 reductase in the ER membrane to the desaturase active site. This type of desaturase also

has been found in plants, such as the sunflower acyl-lipid desaturase (Sperling et al., 1995) and the borage Δ6-fatty acid desaturase (Sayanova et al., 1997).

The DNA sequence differences among the fatty acid desaturase genes is obviously an important factor in the isolation of the various desaturase genes. Due to the conserved amino acid sequenes in the plant fatty acid desaturase superfamily, oligonucleotide probes have been designed in order to screen for other homologous plant genes. For example, the amino acid sequence WVIAHECGH (amino acid residues 101 to 109) of the *Arabidopsis FAD2* protein was used to design an oligonucleotide probe to identify the *Arabidopsis FAD2* gene (Okuley et al., 1994). In this dissertation research, the homologous probe derived from the cDNA of the cotton fatty acid desaturase (*FAD2*) and the heterologous probe from the *Arabidopsis FAD2* gene were successfully used to isolate the cotton *FAD2* gene.

III. Structural analysis and the topology of the cotton ω–6 fatty acid desaturase (FAD2) polypeptide

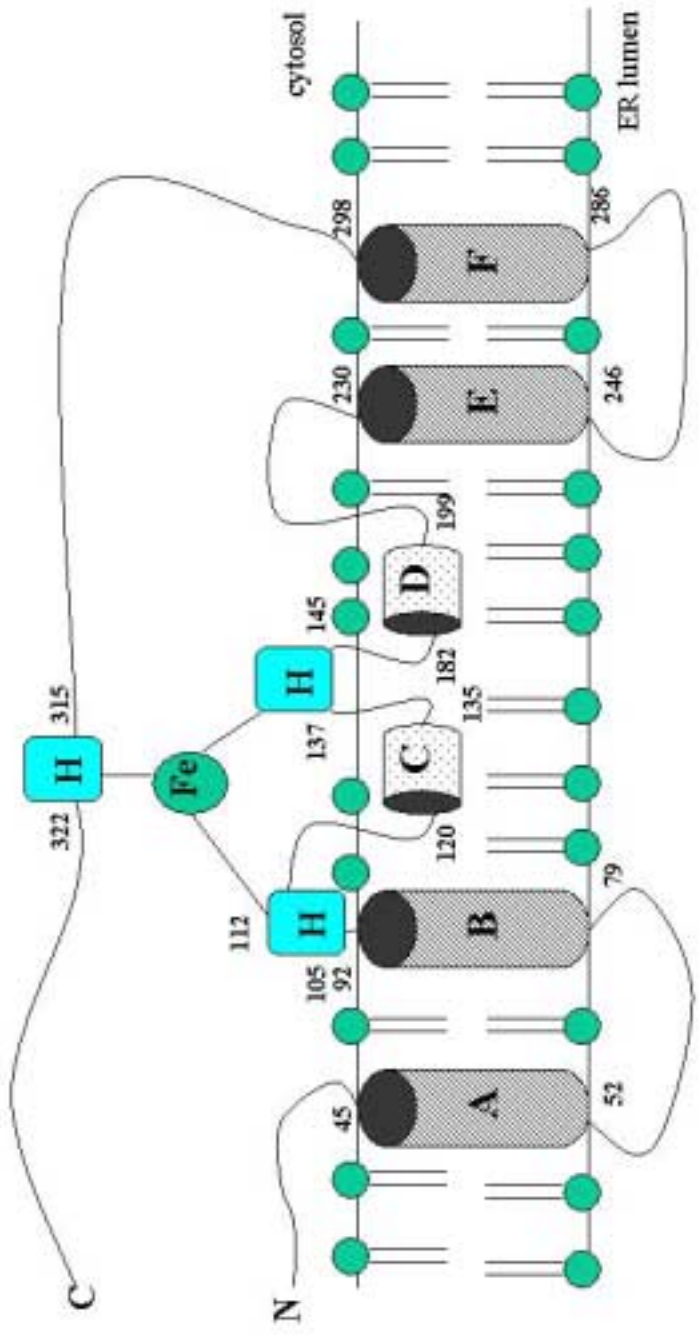
Significant identities occur between the amino acid sequences of the polypeptide products of the cotton *FAD2* gene and other plant ω–6 fatty acid desaturase genes, such as the *G. hirsutum FAD2-1* and *FAD2-2* genes (Liu et al., 1997; 1999), the *G. max FAD2-1* gene, and the *Arabidopsis FAD2* gene. The conserved amino acid residues may be essential in the catalytic activity of these desaturases. This hypothesis leads to predictions for the structure of the cotton *FAD2* polypeptide, as for example, the presence of the diiron-binding motifs. The soluble stearoyl (C18:0) acyl-carrier protein desaturase from *Ricinus* was purified and its three-dimensional structure indicates that it

is a diiron-oxo protein in which the two iron atoms form a bridge with oxygen (Fox et al., 1993). The two iron atoms are coordinated with two histidine and carboxylic acid residues (Fox et al., 1993). Due to the utilization of O_2 and reducing electrons from an electron transport chain, the mechanism of the desaturation reaction of the cotton FAD2 desaturase is probably similar to the catalytic reaction of the soluble stearoyl-18:0 acyl-carrier protein desaturase (Tocher et al., 1998) which has two D/EXXH motifs that provide four diiron-oxo ligands. The amino acid sequence of the cotton FAD2 polypeptide was compared with the protein sequences for seven higher plants and one cyanobacterium (Figure 13) to find the D/EXXH motifs for the iron binding sites. Interestingly, there were three regions of strong conservation among these integral membrane desaturases. One of the most strongly conserved motifs (in the cotton FAD2) was WKYSHRRHH (residues 140 to 148). Another strongly conserved motif occurs in the carboxyl terminus from amino acid residues 315 to 322 (HVAHHLFS). These two conserved motifs were present in all seven higher plant amino acid sequences, except that of the *Synechocystis* Δ12 desaturase, in which the spacing between three histidine residues is increased between residues. However, the FAD2 sequence motif HECCHHF (residues 105 to 112) has homology in all the higher plant proteins, except for the *G. hirsutum* FAD2-1 and (HEXGHHF in FAD2-1) and the cyanobacterial FAD protein (HDCGHRSF). From these results, it appears that the cotton FAD2 and all the FAD2 products from plants and cyanobacteria have great identity in the three histidine rich motifs.

Since the FAD2 enzymes are integral membrane proteins in the endoplasmic reticulum (Los and Murata, 1998; Shanklin and Cahoon, 1998), an interesting question concerns the three-dimensional arrangement of the three histidine motifs in the structures of the desaturases. A hydropathy plot was done by the method of Kyte and Doolittle (Devereux et al., 1984; Kyte and Doolittle, 1982) using DNASIS software (Hitachi) in order to predict the possible membrane-spanning hydrophobic sequences of the FAD2 polypeptide. The predicted hydrophobic regions in the structure of the FAD2 protein shown in Figure 15 is similar to the structures of other membrane-bound desaturases (Stukey et al., 1990; Wada et al., 1990; Reddy et al., 1993; Okuley et al., 1994; Los and Murata, 1998; Shanklin and Cahoon, 1998).

The predicted secondary structure of the FAD2 enzyme was done by Chou and Fasman algorithms in the DNASIS software. There were two long hydrophobic regions of about 45 residues each that extend from residues 45 to 92 and residues 230 to 298 (regions A-B and E-F shown in Figures 15 and 30). Each of the two long stretches form two α-helical membrane-spanning domains in the ER membrane. Two additional short hydrophobic regions (residues 120 to 135 and residues 182 to 199) found in the FAD2 polypeptide may be single-pass membrane-monolayer segments (regions C and D in Figure 15 and Figure 30). The hydropathy plot of the FAD2 enzyme is highly similar to the hydropathy plot of the *Arabidopsis* FAD2 enzyme that also has the additional two potential (single-pass) membrane-monolayer sequences (Okuley et al., 1994). Helix C in Figure 30 probably corresponds to one of the single-pass membrane domains found in

Figure 30. The predicted three-dimensional structure of the putative cotton FAD2 polypeptide (ω-6 fatty acid desaturase) in the endoplasmic reticulum (ER) membrane. The enzyme spans the ER membrane four times (helixes A, B, E, and F, shown as cylinders), and the three histidine clusters (depicted by rectangles labeled H) are on the cytosolic side of the ER. The three conserved histidine-rich motifs are amino acid residues 105 to 112, 137 to 145 and 315 to 322. The two single-pass monolayer domains are α-helixes C and D (shown as cylinders in the top monolayer). The numbers identify amino acid positions in the FAD2 primary structure of 384 amino acids. This diagram was modified from Figure 2 of Los and Murata (1998).


cytosol
ER lumen
N
C
A
B
C
D
E
F
H
H
H
Fe
45
52
92
105
112
79
120
137
135
145
182
199
230
246
286
298
315
322

the hydropathy plots of yeast and rat (Stukey et al., 1990). Hydrophobic regions greater than 50 amino acids long are assumed to be on the cytosolic side of the ER membrane (Stukey et al., 1990). Therefore, it is possible that three conserved histidine-rich motifs (residues 105 to 112, 137 to 143, and 315-322) may be on the cytosolic side of the ER membrane. Since the putative active-site histidines are near the cytosolic site of the ER, it would be easier for these residues to interact directly with is the cytochrome $b_{5,}$ which is the electron donor to the desaturases (Heinz et al., 1993).

IV. Expression of the cotton *FAD2* gene in yeast cells

There are three criteria for determining the cellular localization of ω-6 desaturase enzymes among higher plants. The first criterion is the relative similarity of the deduced amino acid sequences of both the ER and plastid ω-6 fatty acid desaturases (Ohlrogge and Browse, 1995). The second criterion depends on the N-terminal and C-terminal amino acid sequences. Two lysine or arginine residues in the C-terminal regions are found in the ω-6 fatty acid desaturases that occur in the ER membrane. The two lysine residues may be part of the conserved targeting sequence for retention of the protein in the ER (Jackson et al., 1990). In contrast, the plastidial ω-6 fatty acid desaturases require chloroplast transit peptides containing serine and threonine with hydroxyl groups in the N-terminal regions (Teasdale and Jackson, 1996).

The last criterion is the type of biochemical reaction that the chloroplast ω-6 and ER ω-6 fatty acid desaturases catalyze (Heinz et al., 1990; Ohlrogge and Browse, 1995). In general, both chloroplast and ER ω-6 fatty acid desaturases have a common biochemical reaction mechanism. However, they use different lipid substrates and

immediate electron donors. The ER ω-6 desaturases require membrane phosphatidylcholines as the lipid substrates, cytochrome b_5 as the electron donor, and NADH, NADH:cytochrome b_5 oxidoreductase, and oxygen for its activity. The chloroplast ω–6 desaturases use galactolipids as the lipid substrates, reduced ferredoxin and NADPH as electron carriers, and NADPH: ferredoxin oxidoreductase for its activity (Schmidt and Heinz, 1990). From the predicted structure and the hydropathy plot of the cotton ω-6 fatty acid desaturase (FAD2), this enzyme is most likely an integral transmembrane protein with the histidine-conserved motifs.

Expression of the *FAD2* gene was done to study the biochemical reaction and substrate specificity of the cotton FAD2 protein. If the cotton FAD2 protein is indeed an ER ω-6 desaturase, it should function in the presence of an ER membrane environment and a cytochrome b_5 electron donor. Thus, the cotton *FAD2* gene in the plasmid construct pYES2/FAD2 was transformed into *S. cerevisiae* cells, which contains the endogenous substrate oleic acid (C18:1) produced by the yeast *OLE1* gene (Kajiwara et al., 1996; Covello and Reed, 1996). However, these yeast cells cannot convert oleic acid (C18:1) into linoleic acid (C18:2), since they lack any *FAD2* genes (Tocher et al., 1998). The transformed yeast cells with pYES2/FAD2 accumulate linoleic acid (C18:2) which is not normally present in wild-type yeast cells. In addition, a slight amount of palmitoleic acid (C16:2) was also found in the yeast transformants with the pYES2/FAD2 plasmid. The control yeast cells transformed only with the vector pYES2 had no detectable levels of C18:2 and C16:2 fatty acids. Also, the oleic acid (C18:1) peak in the transformed yeast cells was dramatically smaller than the corresponding

C18:1 peak in the control cells. Thus, the dramatic shift in the fatty acid profile of the transformed yeast cells with the cotton *FAD2* gene indicated a substantial increase in linoleic acid and a commensurate decrease in the amount of oleic acid.

The expression of the cotton *FAD2* gene in yeast cells indicates the fidelity and substrate specificity of the FAD2 enzyme. The expression of the *FAD2* gene may depend on temperature, tissue development (Heppard et al., 1996; Kirsch et al., 1997), and gene silencing (Martin et al., 1999). There are many aspects of the *FAD2* gene expression and function that have not been investigated. The evidence shows that cotton microsomal ω-6 desaturase (*ghFAD2-1*) mRNA is induced during embryo development (Liu et al., 1999). The soybean *FAD2-1* gene is expressed in developing seeds (Heppard et al., 1996). The question of whether the expression of the cotton *FAD2* gene is inducible or constitutive still has not been answered. The introduction of the cotton *FAD2* gene into tobacco or *Arabidopsis* plants by *Agrobacterium*-mediated gene transfer may help understand the expression of this gene at the transcriptional and translational levels in plants.

The level of transcription of the soybean *FAD2-1* and *FAD2-2* genes does not increase at cold temperatures (Heppard et.al., 1996). The *Arabidopsis FAD2* gene does not seem to be affected by cold temperatures (Okuley et al., 1994), as compared to the effect of cold temperature on expression of the chloroplast ω-6 fatty acid desaturase (*FAD7*) gene. Therefore, the effect of temperature expression of the cotton *FAD2* gene is still unknown.

Recently, Murakami et al. (2000) provided evidence of thermotolerance in tobacco transgenic plants by silencing the *Arabidopsis* ω-3 fatty acid desaturase (FAD7) in the thylakoid membrane at elevated temperatures. The transgenic tobacco plants with the *Arabidopsis FAD7* gene had induction of unsaturated fatty acids (C18:3) into the thylakoid membranes at 40ºC. The reduction in lipid unsaturation improved the rate of photosynthesis at 40ºC and plant growth at 36ºC (Murakami et al., 2000). Therefore, gene silencing could be one mechanism in controlling the expression of the fatty acid desaturase genes. Moreover, expression of some plant fatty acid desaturase genes may not only be affected by temperature but also by hormones, wounding, or infection (Hamada et al., 1996b; Gadea et al., 1996). Thus, the study of *FAD* gene promoters and the identification of their *tran*-acting factors could help to understand the different mechanisms that regulate the plastidial and microsomal isoforms of ω-6 fatty acid desaturases (Kirsch et al., 1997).

Desaturase gene copy numbers may influence the efficiency of gene expression during temperature change and tissue development (Scheffler et al., 1997). This effect suggests gene knock-out experiments using antisense FAD mRNAs that can inhibit the production of the desaturase gene products (Scheffler et al., 1997). Liu et al. (1999) have reported that there are at least five gene family members for microsomal ω-6 desaturases found in the cotton genome. At least three desaturase genes occur in the A genome of *G. herbaceum* and at least two genes occur in the D genome of *G. raimondii* (Liu et al., 1999). Similarly, multiple gene copies have been reported for stearoyl-ACP desaturase in the *B. napus* genome (Slocombe et al., 1994; Scheffler et al., 1997). The

copy number for this *FAD2* gene was studied by genomic blot bybridization (See Figures 25 and 26). There are two distinct fragments in cotton genomic DNA digested with the restriction enzymes *Hin*dIII, *Hin*dIII/*Sal*I, and *Hpa*I, indicating that there are at least two actual *FAD2* genes in the cotton genome. *G. hirsutum*, as well as *B. napus*, have allotetraploid genomes (Liu et al., 1999). Thus, the A and D genomes of tetraploid cotton have probably diverged from common cotton diploid genome ancestors (Scheffler et al., 1997; Liu et al., 1999).

V. Biotechnology applications with the cotton *FAD2* gene

Recently, great progress has been made producing modified plants using genetic engineering. For example, scientists at DuPont have cloned several genes for soybean enzymes that are important in fatty acid biosynthesis, and have generated transgenic soybean somatic embryos in order to manipulate enzyme activity levels up and down for each step in the complex lipid biosynthesis pathway (Kinney, 1998). It is possible to design a cotton lipid profile to provide a higher quality of cottonseed oil, such as a reduction in harmful saturated fatty acids. In addition, transgenic plants with an increased level of polyunsaturated fatty acids in membrane phospholipids and seed lipids are required to survive wide seasonal variations in growth temperatures and have enhanced tolerance to diseases (Somerville and Browse, 1991; Martin et al., 1999).

The cotton FAD2 enzyme may be commercially important because its decrease in expression in cotton seeds could improve the nutritional value and increase the oxidative stability of cottonseed oil. A new oil with higher amounts of oleic acid (C18:1) can be substituted for oils with harmful levels of saturated fatty acids such as palmitic

acid which enhance the likelihood of coronary heart disease (McNamara, 1992). This is because the consumption of palmitic acid increases levels of LDL-cholesterol and reduces levels of HDL-cholesterol in plasma (McNamara, 1992).

There are many studies describing the modification of plant oils for dietary concerns (Kinney, 1998). A molecular genetics approach can help to manipulate the genes encoding enzymes involved in fatty acid biosynthesis. Much evidence shows that alteration of the expression of a single enzyme activity can change the fatty acid composition in plants (Ohlrogge, 1994). Thus, to increase the content of unsaturated fatty acids like oleic acid in cottonseed oil, the antisense mRNA of both the C16:0-ACP thioesterase (*FatB*) gene and the C18:2-CoA fatty acid desaturase (*FAD2*) gene could be introduced into the cotton genome to produce transgenic cotton plants that have different compositions of edible oils (e.g. with C18:2 fatty acids). Both *FatB* and *FAD2* antisense RNAs could block the flux of polyunsaturated fatty acid biosynthesis passing through the seed compartment rather than through the plastid counterpart (Kinney, 1998). High-oleic acid cotton plants might produce a low-viscosity oil with good oxidative stability for food products and mineral oil lubricants. Also, some functional groups (such as epoxy groups) could be added into oleic acid, which is required for many polyvinylchloride plasticizer applications (Kinney, 1998).

The second application of molecular biology to the study of *FAD2* gene expression is to generate transgenic cotton plants that will be able to survive temperature stress and injury. For example, *FAD2* mutants in *Arabidopsis* showed a dramatic chilling-sensitivity phenotype when grown at 6°C (Miguel et al., 1993). Therefore, the

FAD2 gene is involved in the growth and survival of plants in some manner. The agricultural significance of this chilling phenomenon may lead to the engineering of novel transgenic cotton plants that have an increased content of high melting-point phospholipids to maintain membrane fluidity (Murata et al., 1992; Miquel et al., 1993). Also, Murakami et al. (2000) demonstrated thermotolerance in transgenic plants with the *FAD7* gene when these plants had normal photosynthetic activity at a high temperature (40°C). Thus, genetic engineering of desaturase enzymes could be used to produce temperature-tolerant plants during heat stress (Sharkey, 2000; Murakami et al., 2000). Thus, it is possible to genetically engineer plants with fatty acid desaturase genes for commercial advantage and also for vigorous and vital improvement of plant vigor.

REFERENCE LIST

Aki, T., Shimada, Y., Inagaki, K., Higashimoto, H., Kawamoto, S., Shigeta, S., Ono, K., and Suzuki, O. (1999) Molecular cloning and functional characterization of rat Δ-6 fatty acid desaturase. *Biochem. Biophys. Res. Com*. 255:575-579.

Arondel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H.M., and Somerville, C.R. (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in Arabidopsis. *Science* 258:1353-1355.

Avelange-Macherel, M.H., Macherel, D., Wada, H., and Murata, N. (1995) Site-directed mutagenesis of histidine in the Δ12 acyl-lipid desaturase of *Synechocystis*. *FEBS Lett*. 361:111-114.

Becker, D.M., and Guarente, L. (1991) High-efficiency transformation of yeast by electroporation. Methods in Enzymology, Vol. 194, C. Guthrie and G.R. Fink, eds., Academic Press, New York, NY, pp. 182-187.

Benton, W. D., and Davis, R.W. (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. *Science* 196:180-182.

Browse, J. (1991) Glycerolipid synthesis:biochemistry and regulation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42:467-506.

Browse, J., Kunset, L., Anderson, S., Hugly, S., and Somerville, C. R. (1989) A mutant of *Arabidopsis* deficient in the chloroplast 16:1/18:1 desaturase. *Plant Physiol.* 90:522-529.

Browse, J., McConn, M., James, D., and Miquel, M. (1993) Mutants of *Arabidopsis* deficient in the synthesis of α-linolenate. Biochemical and genetic characterization of the endoplasmic reticulum linoleoyl desaturase. *J. Biol. Chem.* 268:16345-16351.

Chapman, K.D., and Trelease, R.N. (1991). Acquisition of membrane lipids by differentiating glyoxysomes: role of lipid bodies. *J. Cell Biol.* 115:995-1007.

Cheesbrough, T.M. (1989) Changes in the enzymes for fatty acids synthesis and desaturation during acclimation of developing soybean seeds to altered growth temperature. *Plant Physiol*. 90:760-764.

Covello, P. S., and Reed, D. W. (1996) Functional expression of the extraplastidial *Arabidopsis thaliana* oleate desaturase gene (*Fad2*) in *Saccharomyces cerevisiae. Plant Physiol*. 111:223-226.

De Lorgeril, M., Renaud, S., Mamelle, N., Salen, P., Martin, J.L., Monjaud, I., Guidollet, J., Touboul, P., and Delaye, J. (1996) Mediterranean alpha-linolenic acid rich diet in secondary prevention of coronary heart disease. *Lancet.* 343:1454-1459.

Devereux, J., Haeberli, P., and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids. Res*. 12:387-395.

Fan, J., Ranu, R.S., Smith, C., Ruan, C., and Fuller, C.W. (1996) DNA sequencing with [α-^{33}P]-labeled ddNTP terminators: a new approach to DNA sequencing with Thermo Sequenase DNA polymerase. *BioTechniques* 21:1132-1137.

Farmer, E. E. (1994) Fatty acid signaling in plants and their associated microorganisms. *Plant Mol. Biol.* 26: 1423-1437.

Feinberg, A.P., and B.Vogelstein. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem*. 132:6-13.

Ferrante, G., and M. Kates. (1983) Pathways for desaturation of oleoyl chains in *Candida lipolytica. Can. J. Biochem. Cell Biol*. 61:1191-1196.

Fox, B.G., Shanklin, J., Somerville, C., and Munck, E. (1993) Stearoyl-acyl carrier protein Δ9 desaturase from *Ricinus communis* is a diiron-oxo protein. *Proc. Natl. Acad. Sci. USA* 90:2486-2490.

Gadea, J., Mayda, M.E., Conejero, V., and Vera, P. (1996) Characterization of defense-related genes ectopically expressed in viroid-infected tomato plants. *Mol. Plant Microbe. Interact.* 9:409-415.

Gibson, S., Arondel, V., Iba, K., and Somerville, C. (1994) Cloning of a temperature-regulated gene encoding a chloroplast ω-3 desaturase from *Arabidopsis thaliana*. *Plant Physiol*. 106:161516-21.

Gietz, D., St. Jean, A., Woods, R.A., and Schiest, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res*. 20:1425.

Gounaris, K., Mannock, D.A., Sen, A., Brain, A.P.R., Williams, W.P., and Quinn, P.J. (1983) Polyunsaturated fatty acyl residues of galactolipids are involved in the control of bilayer/non-bilayer lipid transitions in higher plants. *Biochim. Biophys. Acta.* 732:229-42.

Green, P.J., Kay, S.A., and Chua, N-H. (1987) Sequence-specific interactions of a pea nuclear factor with light-responsive elements upstream of the *rbcS-3A* gene. *EMBO J*. 6:2543-2549.

Green, P.J., Young, M.H., Cuozzo, N., Kanu-Muratami, Y., Silverstein, P., and Chua, N.H. (1988) Binding sites requirements for pea nuclear protein factor GT-1 correlate with sequences required for light-dependent transcriptional activation of the *rbc-3A* gene. *EMBO J.* 7:4036-4044.

Grula, J.W., Hudspeth, R.L., Hobbs, S.L., and Anderson, D.M. (1995) Organization, inheritance and expression of acetohydroxyacid synthase genes in the cotton allotetraploid *Gossypium hirsutum*. *Plant Mol. Biol*. 28:837-846.

Guilfoyle, T.J. (1997) The structure of plant gene promoters. *Genetic Engineering*. 9:15-47.

Hamada, T., Kodama, H., Nishimura, M., and Iba, K. (1996a) Modification of fatty acid composition by over-expression and antisense-expression of a microsomal omega-3-fatty acid desaturase gene in transgenic tobacco. *Transgenic Research.* 5:115-121.

Hamada, T., Nichiuchi, U., Kodama, H., Nishimura, M., and Iba, K. (1996b) cDNA cloning of a wounding-inducible gene encoding a plastid ω-3 fatty acid desaturase from tobacco. *Plant Cell Physiol*. 37:606-611.

Harwood, J.L. (1997) Plant Lipid Metabolism. *Plant Biochemistry*. (P.M. Dey and J.B. Harborne, eds.), pp. 237-271. Academic Press, New York.

Heinz, E. (1993) Biosynthesis of polyunsaturated fatty acids. In: Moore TS Jr (ed) Lipid metabolism in plants. CRC Press, Boca Raton, Fla., pp 34-89.

Heppard, E.P., Kinney, A.J., Stecca, K.L., and Miao, G. (1996) Developmental and growth temperature regulation of two different microsomal ω-6 desaturase genes in soybeans. *Plant Physiol*. 110:311-319.

Hitz, W.D., Carlson, T.J., Booth, J.R., Kinney, A.J., Stecca, K.L., and Yadav, N.S. (1994) Cloning of a higher-plant plastid ω-6 fatty acid desaturase cDNA and its expression in a cyanobacterium. *Plant Physiol*. 105:635-641.

Iba, K., Gibson, S., Nishiuchi, T., Fuso, T., Nishimura, M., Arondel, V., Hugly, S., and Somerville, C. (1993) A gene encoding a chloroplast ω-3 fatty acid desaturase complements alterations in fatty acid desaturation and chloroplast copy number of the *FAD7* mutant of *Arabidopsis thaliana*. *J. Biol. Chem*. 268:24099-24105.

Itoh. R., Toda, K., Takahashi, H., Takano, H., and Kuroiwa, T. (1998) Delta-9 fatty acid desaturase gene containing a carboxyl-terminal cytochrome b_5 domain from the red alga *Cyanidioschyzon merolae*. *Curr. Genet.* 33:165-170.

Jackson, M.R., Nilsson, T., and Peterson, P.A. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J.* 9:3153-3162.

Jeffcoat, R., Brawn, P.R., Safford, R., and James, A.T. (1977) Properties of rat liver microsomal stearoyl-coenzyme A desaturase. *Biochem. J*. 161:431-437.

Kawagoe, Y., Campbel, B.R., and Murai, N. (1994) Synergism between CACGTG (G-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein beta-phaseolin gene. *Plant J.* 5:885-890.

Kajiwara, S., Shirai, A., Fujii, T., Toguri, T., Nakamura, K., and Ohtaguchi, K. (1996) Polyunsaturated fatty acid biosynthesis in *Saccharomyces cerevisiae* expression of ethanol tolerance and the *FAD2* gene from *Arabidopsis thaliana*. *Appl. Env. Microbiol*. 62:4309-4313.

Kearns, E.V., Hugly, S., and Somerville, C.R. (1991) The role of cytochrome b_5 in Δ12 desaturation of oleic acid by microsomes of safflower (*Cathanus tinctorias L.*). *Arch. Biochem. Biophys*. 284:431-436.

Kinney, A.J. (1998) Plants as industrial chemical factories-new oils from genetically engineered soybeans. *Fett/Lipid* 100:173-176.

Kirsch, C., Hahlbrock, K., and Somssich, I.E. (1997) Rapid and transient induction of a parsley microsomal Δ12 fatty acid desaturase mRNA by fungal elicitor. *Plant Physiol*. 115:283-289.

Kodama, H., Akagi, H., Kusumi, K., Fujimura, T., and Iba, K. (1997) Structure, chromosomal location and expression of a rice gene encoding the microsome ω-3- fatty acid desaturase. *Plant Mol. Biol*. 33:493-502.

Kodama, H., Hamada, T., Horiguchi, G., Nishimura, M., and Iba, K. (1994) Genetic enhancement of cold tolerance by expression of a gene for chloroplast ω-3 fatty acid desaturase in transgenic tobacco. *Plant Physiol*. 105:601-605.

Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.

Lindquist, Y., Huang, W., Schneider, G., and Shanklin, J. (1996) Crystal structure of Δ9 stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins. *EMBO. J.* 15:4081-4092.

Liu, Q., Singh, S.P., Brubaker, C.L., Sharp, P.J., Green, A.G., and Marshall, D.R. (1997) In physiology, Biochemistry and Molecular Biology of Plant Lipids, William, J.P., Khan, M.U., and Lem, N.W., eds. Kluwer Academic Publishers, Boston, MA, pp.383-385.

Liu, Q., Singh, S.P., Brubaker, P.J., Sharp, P.J., Green, A.G., and Marshall, D.R. (1999) Molecular cloning and expression of a cDNA encoding a microsomal ω-6 fatty acid desaturase from cotton (*Gossypium hirsutum*). *Aust. J. Plant Physiol.* 26:101-106.

Los, D.A., and Murata, N. (1998) Structure and expression of fatty acid desaturases. *Biochem. Biophys. Acta.* 1394:3-15.

Martin, M., Leon, J., Dammann, C., Albar, J.P., Griffiths, G., and Sanchez- Serrano, J.J. (1999) Antisense-mediated depletion of potato leaf ω-3 fatty acid desaturase lowers linolenic acid content and reduces gene activation in response to wounding. *Eur. J. Biochem.* 262:283-290.

McNamara, D. (1992) Dietary fatty acids lipoproteins and cardiovascular disease. *Adv. Food Nutr. Res.* 36:254-351.

Meesters, P.A., Springer, J., and Eggink, G. (1997) Cloning and expression of the Δ9 fatty acid desaturase gene from *Cryptococcus curvatus* ATCC20509 containing histidine boxes and a cytochrome b_5 domain. *Appl. Microbiol. Biotechnol.* 47:663-667.

Merion, M., and Warren, W. (1989) Purification of supercoiled plasmids from crude cell lysates using high performance anion exchange chromatography. *BioTechniques.* 7:60-67.

Miquel, M., and Browse, J. (1992) *Arabidopsis* mutants deficient in polyunsaturated fatty acid synthesis. *J. Biol. Chem.* 267:1502-1509.

Miquel, M., James, D Jr, Donner, H., and Browse, J. (1993) *Arabidopsis* requires polyunsaturated lipids for low-temperature survival. *Proc. Natl. Acad. Sci USA* 90:6,208-6,212.

Mitchell, A. G., and Martin, C. E. (1995) A novel cytochrome b_5-like domain is linked to the carboxyl terminus of the *Saccharomyces cerevisiae* Δ9 fatty acid desaturase. *J. Biol. Chem.* 270:29766-29772.

Mitchell, P. J., and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins. *Science* 245:371-378.

Mizutani, M., Tasaka, Tanaka, Y., Ashikuri, T., Kusimi, T., and Murata, N. (1998) Characterization of Δ9 acyl-lipid desaturase homologues from *Arabidopsis thaliana*. *Plant Cell Physiol.* 39:247-253.

Murakami, Y., Tsuyama, M., Kobayashi, Y., Kodama, H., and Iba, K. (2000) Trienoic fatty acids and plant tolerance of high temperature. *Science* 287:476-479.

Murata, N., Ishizaki-Nishizawa, O., Higashi, S., Hayashi, H., Tasaka, Y., and Nishida, J. (1992) Genetically engineered alteration in the chilling sensitivity of plant. *Nature* 356:710-713.

Murata, N., and Wada, H. (1995) Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria. *Biochem. J.* 308:1-8.

Murphy, D.J., Woodrow, I.E., and Mukherjee, K.D. (1985) Substrate specificities of the enzymes of the oleate desaturase system from photosynthetic tissue. *Biochem. J.* 225:267-270.

Napier, J.A., Hey, S.H., Lacey, D.J., and Shewry, P.R. (1998) Identification of a *Caenorhabditis elegans* Δ6-fatty-acid-desaturase by heterologous expression in *Saccharomyces cerevisiae. Biochem. J.* 330:611-614.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., Mcintosh, L., Ohlrogge, J., Raikhel, N., and Somerville, C. (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* 106:1241-1255.

Ni, W., and Trelease, R.N. (1991) Two genes encode the two subunits of cottonseed catalase. *Arch. Biochem. Biophys*. 289:237-243.

Nishiuchi, T., Nakamura, T., Abe, T., Kodama, H., Nishimura, M., and Iba, K. (1995) Tissue-specific and light-responsive regulation of the promoter region of the *Arabidopsis thaliana* chloroplast ω-3 fatty acid desaturase gene (*FAD7*). *Plant Mol. Biol*. 29:599-609.

Ohlrogge, J.B. (1994) Design of new plant products: engineering of fatty acid metabolism. *Plant Physiol.* 104:821-826.

Ohlrogge, J.B., and Browse, J. (1995) Lipid Biosynthesis. *Plant Cell* 7:957-970.

Ohlrogge, J.B., Browse, J., and Somerville, C.R. (1991) The genetics of plant lipids. *Biochim. Biophys. Acta*. 1082:1-26.

Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E., and Browse, J. (1994) *Arabidopsis FAD2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* 6:147-158.

Paterson, A.H., Brubaker, C.L., and Wendel, J.F. (1993) A rapid method for extraction of cotton (*Gossypium spp.)* genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep.* 11:122-127.

Reddy, A.S., Naccio, M.L., Gross, L.M., and Thomas, T.L. (1993) Isolation of a ω-6-desaturase gene from the cyanobacterium *Synechosystis sp.* Strain PCC6803 by gain-of-function expression in *Anabaena sp.* Strain PCC 7120. *Plant Mol. Biol.* 22:293-300.

Sakamoto, T., and Bryant, D.A. (1997) Temperature-regulated mRNA accumulation and stabilization for fatty acid desaturase genes in the cyanobacterium *Synechococcus sp.* Strain PCC 7002. *Mol. Microbiol*. 23:1281-1292.

Sakamoto, T., Los, D.A., Higashi, S., Wada, H., Nishida, J., Ohmori, M., and Murata, N. (1994a) Cloning of ω-3 desaturase from *Cyanobacteria* and its use in altering the degree of membrane-lipid unsaturation. *Plant Mol. Biol.* 26:249-263.

Sakamoto, T., Wada, H., Nishida, I., Ohmori, M., and Murata, N. (1994b) Δ9 Acyl-lipid desaturases of cyanobacteria. Molecular cloning and substrate specificities in terms of fatty acid *sn*-positions, and polar head groups. *J. Biol. Chem*. 269:25576-25580.

Sambrook, J., Fritech, E.F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbor, NY.

Sayanava, O., Smith, M.A., Lapinskas, P., Stobart, A.K., Dobson, G. Christie, W.W., Shewry, P.R., and Napier, J.A. (1997) Expression of a borage desaturase cDNA containing an N-terminal cytochrome b_5 domain results in the accumulation of high levels of Δ6-desaturased fatty acids in transgenic tobacco. *Proc. Natl. Acad. Sci. USA*. 94:4211-4216.

Schmidt, H., Dresselhans, T., Buck, F., and Heinz, E. (1994) Purification and PCR-based cDNA cloning of a plastidial n-6 desaturase. *Plant. Mol. Biol.* 26:631-642.

Scheffler, J.A., Sharpe, A.G., Schmidt, H., Sperling, P., Parkin, I., Luhs, W., Lydiale, D.J., and Heinz, E. (1997) Desaturase multigene families of *Brassica napus* arose through genome duplication. *Theor. Appl. Genet.* 94:583-591.

Schmidt, H., and Heinz, E. (1990) Desaturation of oleoyl groups in envelope membranes from spinach chloroplast. *Proc. Natl. Acad. Sci. USA* 87:9477-9480.

Schneider, G., Lindquivist, Y., Shanklin, J., and Somerville, C. (1992) Preliminary crystallographic data for stearoyl-acyl carrier desaturases from castor seed. *J. Mol. Biol.* 225:561-564.

Shanklin, J., and Cahoon, E.B. (1998) Desaturation and related modifications of fatty acids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:611-641.

Shanklin, J., Whittle, E., and Fox, B.G. (1994) Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* 33:12787-12794.

Sharkey, T.D. (2000) Some like it hot. *Science* 287:43-44.

Slabas, A.R., and Fawcett, T. (1992) The biochemistry and molecular biology of plant lipid biosynthesis. *Plant Mol. Biol.* 19:169-191.

Slocombe, S.P., Piffanelli, P., Fairbairn, D., Bowra, S., Hatzopoula, P., Tsiantis, M., and Murphy, D.J. (1994) Temporal and tissue-specific regulation of a *Brassica napus* stearoyl-acyl carrier protein desaturase gene. *Plant Physiol.* 104:1,167-1,176.

Somerville, C., and Browse, J. (1991) Plant lipids metabolism, mutants, and membranes. *Science* 252:80-87.

Sperling, P., Schmidt, H., and Heinz, E. (1995) A cytochrome b_5-containing fusion protein similar to plant acyl lipid desaturases. *Eur. J. Biochem.* 232:798-805.

Stukey, J.E., McDonough, V. M., and Martin, C.E. (1990) The *OLE1* gene of *Saccharomyces cerevisiae* encodes the delta 9 fatty acid desaturase and can be functionally replaced by the rat stearoyl-CoA desaturase gene. *J. Biol. Chem.* 265:20144-20149.

Taylor, D.C., Barton, D.L., Rionx, K.P., Mackenzie, S.L., Reed, D.W., Underhill, E.W., Pomeroy, M.K., and Weber, N. (1992) Biosynthesis of acyl lipids containing very-long chain fatty acids in microspore-derived and zygotic embryos of *Brassica napus l. cv* Reston. *Plant Physiol*. 99:1609-1618.

Teasdale, R.D., and Jackson, M.R. (1996) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the Golgi apparatus. *Annu. Rev.Cell Dev. Biol.* 12:27-54.

Tiemeier, D.C., Tilghman, and Leder, P. (1977) Purification and cloning of a mouse ribosomal gene fragment in coliphage lambda. *Gene* 2:173-191.

Tocher, D.R., Leaver, M.J., and Hodgson, P.A. (1998) Recent advances in the biochemistry and molecular biology of fatty acyl desaturases. *Prog. Lipid Res.* 37:73-117.

Tosco, A., Gargano, S., Kobayashi, G.S., and Maresca, B. (1997) An AP1 element is involved in transcriptional regulation of Δ9-desaturase gene of *Histoplasma capsulatum. Biochem. Biophys. Res. Commun.* 230:457-461.

Van de Loo, F., and Somerville, C.R. (1994) A plastid omega-3 desaturase from Castor (*Ricinus mumunis L.*). *Plant Physiol*. 105:443-444.

Wada, H., Gombos, Z., and Marata, N. (1990) Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation. *Nature* 347:200-203.

Wada, H., and Murata, N. (1989) *Synechocystis* PCC6803 mutants defective in desaturation of fatty acids. *Plant Cell Physiol*. 30:971-978.

Wang, G., Blakesley, R.W., Berg, D.E., and Berg, C.M. (1993) pDUAL: A transposon-based cosmid cloning vector for generating nested deletions and DNA sequencing templates *in vivo*. *Proc. Natl. Acad. Sci. USA* 90:7874-7878.

Williams, M.E., Foster, R., and Chau, N. (1992) Sequences flanking the hexameric G-BOX Core CACGTG affect the specificity of protein binding. *Plant Cell* 4:485-496.

Yadav, N.S., Wierzbicki, A., Aegerter, M., Caster, C.S., Perez-Grau, L., Kinney, A.J., Hitz, W.D., Booth, J.R., Carlson, T.J., Russell, S.H., Feldmann, K.A., Pierce, J., and Browse, J. (1993) Cloning of higher plant ω-3 fatty acid desaturases. *Plant Physiol.* 103:467-476.

York, D., Welch, K., Goryshin, I.Y., and Reznikoff, W.S. (1998) Simple and efficient generation *in vitro* of nested deletions and inversions: *Tn5* intramolecular transposition. *Nucl. Acids Res.* 26:1927-1933.

Young, R.A. (1991) RNA polymerase II. *Annu. Rev. Biochem*. 60:689-715.